Conducting and reporting case series and audits
– author guidelines for Acupuncture in Medicine

Adrian White

Abstract
This article aims to give guidance on the conduct and reporting of case series and audits of acupuncture, based on common problems that have arisen in the past. This type of project, and particularly the prospective case series or pre-post-intervention study, may give valuable evidence of the overall effectiveness of acupuncture – for example in different situations and in different conditions – and provides one step in the research pathway before generating an hypothesis. The project should be designed with the aim of reducing bias as much as possible. Careful and detailed planning is essential for the project to produce worthwhile results that readers can evaluate and replicate. Ethical issues should be considered and formal approval may be necessary. The patient group should be recruited systematically and baseline data obtained. The treatment given should be systematic and decisions to change or end treatment made explicit. The outcome should be measured in ways that are known to be reliable and valid. Musculoskeletal problems can be evaluated with scales for pain and confirmed by measuring one other symptom such as stiffness or one other aspect such as bothersomeness. Global change scores also provide supporting information, and the MYMOP (Measure Yourself Medical Outcome Profile) measure is popular when patients with different conditions are included. The design of questionnaires for beliefs and attitudes is a specialised area that should not be attempted without expert help. Adverse events should also be recorded. Analysis of the data and the best way of summarising and presenting the results are also discussed.

Introduction
We often receive submissions of articles that describe how patients have responded to a course of acupuncture treatment. Some of these articles are unsatisfactory for a variety of reasons, and may only be retrieved with a great deal of hard work by the authors and the editorial team, or they may even be beyond rescue and have to be rejected, a great disappointment for the author. This article gives guidelines that are based on the experience of reviewing previous reports and aims to guide authors on how to plan, conduct and report these studies in a way that will be readily publishable.

Acupuncture practitioners do this kind of study for various reasons: for example, they might want to examine the effect of a particular form of acupuncture; the effect of acupuncture on a particular type of patient or particular conditions they are interested in; the factors that predict which patients respond to treatment and which do not; or the results of treatment in a particular setting. Case series can be used to estimate costs of treatment but are less useful for this than controlled trials. They may even simply evaluate routine treatment in routine patients, though journal editors and readers will be more interested in studies that do something different. One other motive for conducting a case series is personal development, since it provides an excellent experience in planning, conducting, analysing and reporting a piece of research (but beware of the ethics of this, see below).

A fundamental mark of good quality in this kind of study is when patients have been assessed before and after treatment. This usually means setting up a prospective case series (see Table 1) sometimes called a pre-post-intervention study. Occasionally, a retrospective audit can achieve this, if symptoms have been routinely scored at the beginning and end of treatment. It is important to realise what conclusions you can legitimately draw from this kind of observational study: it can provide evidence of the overall effectiveness of acupuncture treatment; but you cannot conclude that it was the needling that caused the improvement – it may have been expectation, time or various other influences. To
produce evidence of cause and effect requires an experimental study in which variables are manipulated and groups are compared. Observational studies can be used to influence colleagues or to develop the service itself, for example by improving referrals. Also, and importantly, case series provide the second stage of evidence in a stepwise research pathway: after case reports and anecdotes have described a phenomenon, case series can quantify it and provide the background for a hypothesis that can then be tested using a controlled trial.

In case series, like any other project, it is important to reduce the effect of the natural bias of both the researcher and the patients. There are four main categories of bias. 1) Selection bias, for example excluding patients who you thought would not respond to acupuncture, so they do not spoil the overall picture. 2) Measurement bias, for example phrasing the questions in such a way that the answer is likely to be positive; retrospective studies are usually fatally flawed by measurement bias: patients give glowing reports of the outcome, partly because they do not remember how severe their symptoms were, and partly so as not to offend their doctor (as in, “They would say that, wouldn’t they?”). 3) Attrition bias, for example not making allowance for those patients who drop out of the study: they are least likely to have responded to treatment, so omitting their results increases the apparent overall effect. 4) Performance bias, which only applies to controlled trials: the groups receive some systematic difference in care or measurement, other than the intervention being tested.

Acupuncturists should recognise their own bias which comes with their enthusiasm for the therapy. It can reduce the objectivity of the project and make the results fairly meaningless. The medical profession makes a virtue of criticising bias in pharmaceutical research, and should apply the same rules to acupuncture research: bias interferes with the best use of healthcare resources and is against the patients’ interests.

One final introductory point about case series is that they are little use if the reader cannot understand what was done and make a critical evaluation of the results. For this reason, studies should be of a high standard, and the patients, the treatment and the outcomes should be described accurately.

Planning
It is essential to think through every detail of the study in advance – patients, measures, data collection, analysis and reporting – and have your plan reviewed by a colleague or an established researcher. Although ‘plan it thoroughly’ sounds like a platitude, lack of planning has been conspicuous in more than one article submitted to the journal. The value of arranging for a peer to review your plans cannot be overemphasised.

Planning starts by clearly identifying why you are doing this particular case series, as there are often mixed motives – commonly, to justify the time spent giving acupuncture and to count towards professional certification. Overt attempts to score ‘political’ points with health trusts, for example, may sit awkwardly

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Formal definitions of some forms of observational study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>Measurement*</td>
</tr>
<tr>
<td>Case series</td>
<td>Prospective or retrospective</td>
</tr>
<tr>
<td>Cohort study</td>
<td>Prospective</td>
</tr>
<tr>
<td>Audit</td>
<td>Retrospective</td>
</tr>
</tbody>
</table>

*Prospective - study is initiated before the outcome is measured. Retrospective - study is initiated after the outcome is measured.
with the need to be objective, perhaps by influencing the wording of the questions.

The best case series have been set up after explicit planning meetings which involve the researcher, the staff who are going to be involved with the data collection, and at least one peer who is not directly involved but can comment on the practicality of the plans. A patient representative could also be involved. Another thing that is essential to do in the planning stage is to read previous published case series to cross-check their methods and learn from their mistakes (hopefully described in the Discussion section), and see what the finished report may look like.31

Your plans should cover, in detail, every aspect of the study: how the patients are to be identified and informed, what baseline and outcome information you will collect and precisely how and when, what your treatment will consist of and how standardised it will be, and how you will approach the analysis.

**Ethics and governance**

All research proposals involving human subjects, their records, their tissue, or their carers will need ethical approval, unless the project is defined as audit or evaluation. You should consider the ethical principles involved in your study, anyway: patients need to be asked for their consent, firstly for any extra commitment that they are asked to give (eg completing extra forms, attending extra follow up sessions) and secondly for their clinical information to be used for any reason other than treatment (which also applies to purely retrospective audits). To provide informed consent, patients must have sufficient information to know what the purpose of the study is and how much it intrudes on them personally, and how much it intrudes on them personally, and sufficient time to absorb the information and reflect on it – usually regarded as 24 hours. Written consent is best. If you work in the NHS, then you should discuss the project in a phone-call with the co-ordinator of your Local Research Ethics Committee (LREC) located through the internet. The decision about whether formal approval is required from the ethics committee will probably depend largely on the intrusiveness: one or two questions about pain and stiffness in musculoskeletal conditions is hardly intrusive, but formal approval may well be necessary if patients are asked to complete questionnaires on psychological or obstetric issues, or if they are asked to attend extra sessions. If you work entirely privately, then you need to consider the ethical issues yourself, and possibly find an independent opinion eg from your local academic centre.

It is not sufficient simply to classify your project in your own mind as ‘audit’ or ‘evaluation’ and therefore conclude that ethics approval is not required.

If you are employed by an NHS trust, you will need to discuss issues of research governance with them, ie whether the project has consequences for the delivery of other services, eg through the use of your time or resources used to retrieve the notes. In most cases a phone call to the trust R&D will confirm whether or not there is a need for formal approval.

**Patients**

The way you recruit and select patients for the study will depend on your objective; for example you may want to investigate patients with shoulder pain, patients taking medication, patients over 80 years old, patients who are waiting for knee replacement surgery. You must plan who you will include and exclude, because you must include every patient who is eligible to avoid selection bias. Will you, for example, include patients who cannot complete the outcome measure at baseline? Or exclude those who have only had the condition for two weeks? Will you include those whom you have already treated for the condition, or for some other condition? You should also plan the criteria for stopping treatment.

You would normally recruit patients either over a fixed period of time or (better) until you have recruited the number you need. You would not usually need to do a formal sample size calculation, though you could: it would depend on how big a change you think would be worth demonstrating, and what the mean and standard deviation are at baseline, which you can easily collect from the first half dozen patients. A study with fewer than two dozen patients might not be very interesting for readers, unless the condition was really quite rare.

You will need to describe the patients and their characteristics in your report, which usually means a minimum of their age (mean and standard deviation, or range maximum to minimum), sex distribution, duration and severity of the condition, current use of medication, and any previous treatment for the
Table 2 Topics to consider in reporting a case series

<table>
<thead>
<tr>
<th>Section heading</th>
<th>Subheading, if needed</th>
<th>Outline of contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td></td>
<td>Background to the study and its aims</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Setting including experience of acupuncture</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Style of acupuncture</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ethics and governance issues</td>
</tr>
<tr>
<td>Methods</td>
<td>Patients</td>
<td>Source and selection criteria</td>
</tr>
<tr>
<td></td>
<td>Outcome measurement</td>
<td>Description of measures and reference to validation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Development and testing of any novel questions or measures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverse events</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Details of administration and collection</td>
</tr>
<tr>
<td></td>
<td>Analysis</td>
<td>Measurement points</td>
</tr>
<tr>
<td>Results</td>
<td>Patients</td>
<td>Approach to quantitative data (usually descriptive statistics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qualitative data handling, analysis, validating</td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>Patient flow, reasons for dropouts or withdrawals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Description of patients (table, perhaps bar chart)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Summary findings in text, details in table and figure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Resume of findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Strengths and limitations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparison with other published findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speculation about mechanism, if relevant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Implications for practice and research</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conclusion, if appropriate</td>
</tr>
</tbody>
</table>

condition. How are you going to collect all this information?

You will also, importantly, have to describe the ‘flow of patients’ through the study, and account for any who drop out (their decision) or are withdrawn (your decision) before the end and describe the reasons.

The acupuncture treatment

For the work to be of any use to anyone else, the acupuncture treatment and how you make the decision to change or end treatment needs to be described carefully and accurately, so that it can be reproduced. It is no good reporting that ‘some patients were given electroacupuncture’, for example, as the reader will want to know how you made the decision to use electroacupuncture. All reports of acupuncture treatment should conform to the STRICTA criteria.¹

There are other general points about evaluating patients. You should use the measure in the way that it is designed to be used; the instructions should state the ‘time frame’, for example the average or worst pain over the last 24 hours or the last week, or the present pain; any questions that you devise should allow the patients a full range of responses from ‘very much worse’ to ‘very much better’. A common error among acupuncturists is not to offer patients the option of responding ‘worse’ – this may reflect well on the expectations of the effect of acupuncture, but is still poor science. You should describe how you collected all data, including from patient records if appropriate.

Offering boxes for free text responses may be user friendly but is not researcher friendly: written responses can be difficult to analyse in any meaningful way.

Evaluating the outcome

The point of the exercise is to show the effect of treatment, i.e. the difference in symptoms before and after treatment. You must measure patients in a way that is both reliable – performs the same way with different patients, for example – and valid – measures what it is meant to measure. This effectively means using a measure that has already been developed and tested.

Pain

Since acupuncture is most commonly used for musculoskeletal conditions, many case series will measure pain. Use a simple measure like the visual analogue scale (VAS) or numerical rating scale (NRS) as shown in the Box. Note that 1) no pain is usually zero not one, to avoid confusion; and 2) the right hand end is not ‘worst pain I have experienced’ because that does not leave room to score pain that is
even worse than that. Patients can usually understand either system very easily after an explanation, perhaps using hunger as an example. It is best to try to gain confirmatory evidence about the patients’ response with a second question. This could be about limitation of movement, stiffness, or interference with daily activity, or bothersomeness, depending on the condition. The VAS or NRS may have been validated for this symptom too, or you may have to search for some other validated way of assessing it.

Global change score
It is simple (and therefore common) to ask patients to score their improvement retrospectively, either on a scale or by ticking a box from ‘very much worse’ to ‘very much better’. This measure has limitations: patients cannot usually remember their condition at the moment when they actually started treatment. Any improvement that they record is likely to include the effects of time and other factors too.

### Table 3 Some examples of problems in question design

<table>
<thead>
<tr>
<th>Question</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>What was the attitude of the doctor?</td>
<td>Ambiguous: attitude to what?</td>
</tr>
<tr>
<td>How much of your benefit do you feel was due to acupuncture?</td>
<td>Requires insight into placebo effects</td>
</tr>
<tr>
<td>Do you feel you had the right number of treatments?</td>
<td>Requires insight into the response to acupuncture</td>
</tr>
<tr>
<td>When you started treatment, what was your impression of acupuncture?</td>
<td>Ambiguous, and open to recall bias</td>
</tr>
<tr>
<td>Should this acupuncture clinic continue to be funded?</td>
<td>Strong bias towards positive response</td>
</tr>
<tr>
<td>Has your level of activity changed since the end of treatment?</td>
<td>Ambiguous: changed because of treatment, or for other reasons</td>
</tr>
</tbody>
</table>

Other symptoms
The MYMOP is popular because it is patient-centred, but it was really designed as a single method of assessment of a variety of presenting complaints, from depression through fatigue to joint pain. It is not really applicable if all the patients have similar painful conditions. The instructions for use should be followed carefully, and are available on the website.'

Outcome measures for several other symptoms and conditions like depression and anxiety can be downloaded from the internet, but consider first whether they are reliable and valid – and perhaps ask for advice. It would be very useful to have a simple and reliable method for measuring use of medication such as analgesics, but such a thing does not exist because patients’ recall is not at all accurate.

Beliefs and attitudes
Measuring people’s beliefs and attitudes properly and in a valid way is a highly skilled technique, involving preliminary work to identify dimensions, development and piloting of questionnaires. This is beyond the scope of this article, and not advised without help from an expert, probably someone trained in social science or in psychology. Some examples of questions that illustrate problems are given in Table 3.

Adverse events
Finally, you should routinely ask about the adverse effects of treatment, in order both to collect important information and to demonstrate your fair and balanced approach to the subject of acupuncture. The question should be quite a general one, such as:

---

**Numerical Rating Scale**
Please score your pain from zero to ten, where zero means no pain and ten means the worst pain you can imagine. Please write the score here.

**Visual Analogue Scale**

```
0  100
None  Worst I can imagine
```

---

**Table 3 Some examples of problems in question design**

<table>
<thead>
<tr>
<th>Question</th>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>What was the attitude of the doctor?</td>
<td>Ambiguous: attitude to what?</td>
</tr>
<tr>
<td>How much of your benefit do you feel was due to acupuncture?</td>
<td>Requires insight into placebo effects</td>
</tr>
<tr>
<td>Do you feel you had the right number of treatments?</td>
<td>Requires insight into the response to acupuncture</td>
</tr>
<tr>
<td>When you started treatment, what was your impression of acupuncture?</td>
<td>Ambiguous, and open to recall bias</td>
</tr>
<tr>
<td>Should this acupuncture clinic continue to be funded?</td>
<td>Strong bias towards positive response</td>
</tr>
<tr>
<td>Has your level of activity changed since the end of treatment?</td>
<td>Ambiguous: changed because of treatment, or for other reasons</td>
</tr>
</tbody>
</table>
Have you noticed anything new or unusual which you think could be due to acupuncture? In particular, make it clear in the report whether or not any of the patients dropped out of treatment because of adverse effects.

Outcome measures for explanatory trials

There is no need to give in to the idea that you have to use identical measures to those used in placebo-controlled RCTs that are further up the research ladder. They have different needs, and collect more data for a variety of reasons. Typically, ‘causal’ studies like RCTs will go to the trouble of evaluating the condition in several ways: its symptoms, and its effect on daily function and its effect on quality of life. This is all some considerable effort – both for the patient who has to fill in all the forms, and for the researcher. It all has to be analysed and reported – collecting data and then not using it is unethical.

Data collection and analysis

The report must include details of how you administered the measures for the various data – the baseline information, and the pre- and post-treatment measures. This is because the way you collected them can easily affect their validity. For example, if you give the patients the forms to complete right in front of you, or to hand directly back to you, you will increase the problem of response (measurement) bias: most patients are grateful for your time and effort, and will therefore be inclined to rate the improvement higher than they might otherwise. Similarly, measures administered immediately before the first treatment and immediately after the last could show improvement that is just due to temporary release of opioid peptides!

Quantitative data should be analysed so the report can describe the subjects’ characteristics, details of treatment such as the number of sessions attended and the adverse events, and patients’ symptoms before and after. It is best to report the outcome in the actual form that it was scored, and not to convert the results into arbitrary categories, such as ‘small response’, ‘good response’ and so on. However, it is acceptable to report the numbers of patients who have a 50% improvement in symptoms, in addition to the mean scores. This sort of analysis does not require elaborate statistics software, but can be done with functions of popular spreadsheets such as ‘Excel’. There is a standard way to summarise data, by giving two pieces of information – the average value, and the degree of spread of values. These usually are the mean and standard deviation (SD) respectively. But sometimes it is necessary to report the median and interquartile range (ie 25% to 75%) because the mean and SD do not truly reflect the data, often because they are skewed to one side or other. Ask for help on this point.

Percentages help the reader to understand and interpret the data. They are not part of the precise record, and any decimal place should usually be rounded to the nearest whole number. Percentages of groups of less than 20 are not reliable.

It is usually very complex to analyse qualitative data meaningfully, and likely to demand more expertise than simple quantitative data analysis. It is beyond the scope of this article.

Reporting the project

Before you start to write, look at previous similar papers. Do not hesitate to ask for advice or help from someone experienced in writing reports, or the journal editor. Beware of writing in an obviously biased attitude: you are entitled to feel good about having reached this stage in your study (many studies are performed but never reported) and to make the best of the description, but you should also report the errors and deficiencies, so that readers can both evaluate the strength of your findings, and learn from your mistakes. The standards required for scientific reports are probably higher than for internal reports for employers. It is important to put yourself in the position of the reader and think what information is ‘necessary and sufficient’ for the report. Readers should not have to assume anything – but neither should they expect to be told anything twice.

The article title should indicate the type of paper, eg case series. The abstract should state the numbers of patients and summarise the actual findings, with data. It is always good to have a catchy, memorable ‘sound-bite’ summary of the research somewhere in the report, and the abstract is a good place because it is seen by most people: something like: ‘49% of the patients who completed treatment had maintained benefit after six months’.¹

Report the exact wording of any questionnaire (unless it is a very common one that all readers
are likely to know) and the available responses, preferably in the results section. Giving a reference is not sufficient because many readers will not have access to that article. The reason to report the actual wording is that it can influence the response.

Tables and figures
When presenting the results, the main data should be reported in detail in a table (which by definition has more than one row and more than one column) or in a figure (which is anything that is not a table). Tables have titles but figures have legends, for some reason, so that the words accompanying a figure need to make sense grammatically. Tables and figures should be able to stand independently from the text, though occasionally it is necessary to state ‘see text for further details’.

You should help readers to assimilate the information as quickly as possible. Bar charts are often the easiest way to grasp the message, but they are not good for describing the spread of the data – adding whiskers makes them confusing. This journal is unusual in being willing to have important data represented both in table (for full details including SDs) and in figure form such as bar chart, for quick understanding. Authors should not put the full details of the data in the text as well, but should use it to draw out or summarise the most important results.

To represent the distribution of age or conditions among the group of patients, it is tempting and technically correct to use a pie-chart. However, pie-charts in print are often unattractive and difficult to digest and it is usually best to use a standard bar chart.

Discussion
It is important not to make excessive claims for the interpretation of case series. Remember that patients could have improved because of time or expectation, or other factors. A useful phrase is that acupuncture was ‘associated with’ the improvement.

Style
There is a strong tendency for people who can hold a perfectly normal conversation in English to think that they must become formal and obscure as soon as they start to write a scientific report. This probably happens because research is supposed to be impersonal and objective, and the language somehow needs to reflect this. We encourage authors to use simple, direct language as much as possible and to avoid being too formal. However, it is still necessary to use accurate and precise terms, so that they are not open to different interpretations, and to write succinctly, to save space and time.

Conclusion
Case series can have various aims including documenting the improvements following treatment, searching for prognostic factors, and providing the practitioner with experience in a scientific project. They should be conducted rigorously in planning, performance and report stages. They are straightforward in theory, but they are not necessarily simple; they require some expertise, and common reasons for failing to achieve their objectives are lack of planning, and not recognising the need for help in the early stages.

Acknowledgement
Penny Gardiner provided valuable advice on the ethics and research governance sections.

Reference List
5. MYMOP, Measure Yourself Medical Outcome Profile. http://www.hsrc.ac.uk/mymop/main.htm
Conducting and reporting case series and audits - author guidelines for acupuncture in medicine
Adrian White

doi: 10.1136/aim.23.4.181

Updated information and services can be found at:
http://aim.bmj.com/content/23/4/181

These include:

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/