Understanding your study group is key to getting a good response to a questionnaire; dealing with the resulting mass of data is another challenge.

The first step in producing good questionnaire research is getting the right questionnaire. However, even the best questionnaire will not get adequate results if it is not used properly. This article outlines how to pilot your questionnaire, distribute and administer it; and get it returned, analysed, and written up for publication. It is intended to supplement published guidance on questionnaire research, three quarters of which focuses on content and design.

Piloting

Questionnaires tend to fail because participants don’t understand them, can’t complete them, get bored or offended by them, or dislike how they look. Although friends and colleagues can help check spelling, grammar, and layout, they cannot reliably predict the emotional reactions or comprehension difficulties of other groups. Whether you have constructed your own questionnaire or are using an existing instrument, always pilot it on participants who are representative of your definitive sample. You need to build in protected time for this phase and get approval from an ethics committee.

During piloting, take detailed notes on how participants react to both the general format of your instrument and the specific questions. How long do people take to complete it? Do any questions need to be repeated or explained? How do participants indicate that they have arrived at an answer? Do they show confusion or surprise at a particular response—if so, why? Short, abrupt questions may unintentionally provoke short, abrupt answers. Piloting will provide a guide for rephrasing questions to invite a richer response (box 1).

Planning data collection

You should be aware of the relevant data protection legislation (for United Kingdom see www.informationcommissioner.gov.uk) and ensure that you follow internal codes of practice for your institution—for example, obtaining and completing a form from your data protection officer. Do not include names, addresses, or other identifying markers within your electronic database, except for a participant number linked to a securely kept manual file.

The piloting phase should include planning and testing a strategy for getting your questionnaire out and back—for example, who you have invited to complete it (the sampling frame), who has agreed to do so (the response rate), who you’ve had usable returns from (the completion rate), and whether and when you needed to send a reminder letter. If you are employing researchers to deliver and collect the questionnaire it’s important that they know exactly how to do this.

Administrative errors can hamper the progress of your research. Real examples include researchers giving the questionnaire to wrong participants (for example, a questionnaire aimed at men given to women); incomplete instructions on how to fill in the questionnaire (for example, participants did not know whether to tick one or several items); postal surveys in which the questionnaire was missing from the envelope; and a study of over 3000 participants in which the questionnaire was sent out with no return address.

Box 1: Patient preference is preferable

I worked on a sexual health study where we initially planned to present the questionnaire on a computer, since we had read people were supposedly more comfortable “talking” to a computer. Although this seemed to be the case in practices with middle class patients, we struggled to recruit in practices where participants were less familiar with computers. Their reasons for refusal were not linked to the topic of the research, but because they saw our laptops as something they might break, could make them look foolish, or would feed directly to the internet (which was inextricably linked to computers in some people’s minds). We found offering a choice between completing the questionnaire on paper or the laptop computer greatly increased response rates.
Administering your questionnaire

The choice of how to administer a questionnaire is too often made on convenience or cost grounds (see table A on bmj.com). Scientific and ethical considerations should include:

- The needs and preferences of participants, who should understand what is required of them; remain interested and cooperative throughout completion; be asked the right questions and have their responses recorded accurately; and receive appropriate support during and after completing the questionnaire.
- The skills and resources available to your research team.
- The nature of your study—for example, short term feasibility projects, clinical trials, or large scale surveys.

Maximising your response rate

Sending out hundreds of questionnaires is a thankless task, and it is sometimes hard to pay attention to the many minor details that combine to raise response and completion rates. Extensive evidence exists on best practice (box 2), and principal investigators should ensure that they provide their staff with the necessary time and resources to follow it. Note, however, that it is better to collect fewer questionnaires with good quality responses than high numbers of questionnaires that are inaccurate or incomplete. The third article in this series discusses how to maximise response rates from groups that are hard to research.15

Accounting for those who refuse to participate

Survey research tends to focus on people who have completed the study. Yet those who don’t participate are equally important scientifically, and their details should also be recorded (remember to seek ethical approval for this).10

Box 2: Factors shown to increase response rates

- The questionnaire is clearly designed and has a simple layout.
- It offers participants incentives or prizes in return for completion.
- It has been thoroughly piloted and tested.
- Participants are notified about the study in advance with a personalised invitation.
- The aim of study and means of completing the questionnaire are clearly explained.
- A researcher is available to answer questions and collect the completed questionnaire.
- If using a postal questionnaire, a stamped addressed envelope is included.
- The participant feels they are a stakeholder in the study.
- Questions are phrased in a way that holds the participant’s attention.
- Questionnaire has clear focus and purpose and is kept concise.
- The questionnaire is appealing to look at, as is the researcher.
- If appropriate, the questionnaire is delivered electronically.

One way of reducing refusal and non-completion rates is to set strict exclusion criteria at the start of your research. For example, for practical reasons many studies exclude participants who are unable to read or write in the language of the questionnaire and those with certain physical and mental disabilities that might interfere with their ability to give informed consent, cooperate with the researcher, or understand the questions asked. However, research that systematically excludes hard to reach groups is increasingly seen as unethical, and you may need to build additional strategies and resources into your study protocol at the outset.15 Keep a record of all participants that fit the different exclusion categories (see bmj.com).

Collecting data on non-participants will also allow you to monitor the research process. For example, you may find that certain researchers seem to have a higher proportion of participants refusing, and if so you should work with those individuals to improve the way they introduce the research or seek consent. In addition, if early refusals are found to be unusually high, you might need to rethink your overall approach.10

Entering, checking, and cleaning data

Novice researchers often assume that once they have selected, designed, and distributed their questionnaire, their work is largely complete. In reality, entering, checking, and cleaning the data account for much of the workload. Some principles for keeping quantitative data clean are listed on bmj.com.

Even if a specialist team sets up the database(s), all researchers should be taught how to enter, clean, code, and back up the data, and the system for doing this should be universally agreed and understood. Agree on the statistical package you wish to use (such as SPSS, Stata, Epilinfo, Excel, or Access) and decide on a coding system before anyone starts work on the dataset.

It is good practice to enter data into an electronic database as the study progresses rather than face a mountain of processing at the end. The project manager should normally take responsibility for coordinating and overseeing this process and for ensuring that all researchers know what their role is with data management. These and other management tasks are time consuming and must be built into the study protocol and budget. Include data entry and coding in any pilot study to get an estimate of the time required and potential problems to troubleshoot.

Analysing your data

You should be able to predict the type of analysis required for your different questionnaire items at the planning stage of your study by considering the structure of each item and the likely distribution of responses (box 3). Table B on bmj.com shows some examples of data analysis methods for different types of responses.

Writing up and reporting

Once you have completed your data analysis, you will need to think creatively about the clearest and most
Examples of ways of presenting data and when to use them

<table>
<thead>
<tr>
<th>Type</th>
<th>When to use</th>
<th>When to avoid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data table</td>
<td>If you need to produce something that is simple and quick and that has a low</td>
<td>Do not use if you want to make your work look visually appealing. Too many</td>
</tr>
<tr>
<td></td>
<td>publication cost for journals. If you want to make data accessible to the</td>
<td>tables can weigh down the results section and obscure the really key results.</td>
</tr>
<tr>
<td></td>
<td>interested reader for further manipulations</td>
<td>The reader is forced to work too hard and may give up reading your report.</td>
</tr>
<tr>
<td>Bar chart</td>
<td>If you need to convey changes and differences, particularly between groups</td>
<td>If your data are linear and each item is related to the previous then you</td>
</tr>
<tr>
<td></td>
<td>(eg how men and women differed in their views on an exercise programme for</td>
<td>should use a (line) graph. Bar charts treat data as though they are separate</td>
</tr>
<tr>
<td></td>
<td>recovering heart attack patients)</td>
<td>groups not continuous variables.</td>
</tr>
<tr>
<td>Scatter</td>
<td>Mostly used for displaying correlations or regressions (eg association</td>
<td>If your data are based on groups or aggregated outcomes rather than</td>
</tr>
<tr>
<td>graph</td>
<td>between number of cigarettes smoked and reduced lung capacity)</td>
<td>individual scores.</td>
</tr>
<tr>
<td>Pie chart</td>
<td>Used for simple summaries of data, particularly if a small number of choices</td>
<td>As with bar charts, avoid if you want to present linear or relational data.</td>
</tr>
<tr>
<td></td>
<td>were provided</td>
<td></td>
</tr>
<tr>
<td>Line graph</td>
<td>Where the points on the graph are logically linked, usually in time (eg scores</td>
<td>If your data were not linked over time, repetition, etc it is inappropriate</td>
</tr>
<tr>
<td></td>
<td>on quality of life and emotional wellbeing measures taken monthly over six</td>
<td>to suggest a linear relation by presenting findings in this format.</td>
</tr>
<tr>
<td></td>
<td>months)</td>
<td></td>
</tr>
</tbody>
</table>
Summary points

Piloting is essential to check the questionnaire works in the study group and identify administrative and analytical problems.

The method of administration should be determined by scientific considerations not just costs.

Entering, checking, and cleaning data should be done as the study progresses.

Don’t try to include all the results when reporting studies.

Do include exclusion criteria and data on non-respondents.

The research participant may be unaware of the purpose of the study or skill deficits will make the entire study unsound.

Research participants, on the other hand, may be motivated to complete a questionnaire through interest, boredom, a desire to help others (particularly true in health studies), because they feel pressurised to do so, through loneliness, or for an unconscious ulterior motive ("pleasing the doctor"). All of these introduce potential biases into the recruitment and data collection process.

I thank Alicia O’Cathain, Trish Greenhalgh, Jill Russell, Geoff Wong, Marcia Rigby, Sara Shaw, Fraser Macfarlane, and Will Callaghan for their helpful feedback on earlier versions of this paper and Gary Wood for advice on statistics and analysis.

PMB has taught research methods in a primary care setting for the past 13 years, specializing in practical approaches and using the experiences and concerns of researchers and participants as the basis of learning. This series of papers arose directly from questions asked about real questionnaire studies.

To address these questions she and Trisha Greenhalgh explored a wide range of sources from the psychological and health services research literature.

Competing interests: None declared.

1 Boynton PM, Greenhalgh T. Hands-on guide to questionnaire research: selecting, designing, and developing your questionnaire. BMJ 2004; 328:1312-5.
15 Boynton PM, Wood GW, Greenhalgh T. Hands-on guide to questionnaire research: reaching beyond the white middle classes. BMJ (in press).

(Accepted 17 March 2004)
Acquired haemophilia A may be associated with clopidogrel
Montaser Haj, H Dasani, S Kundu, U Mohite, P W Collins

Acquired haemophilia A is a rare bleeding disorder caused by autoantibodies against factor VIII. Bleeding is often severe and may be life threatening. In half of patients, no underlying disorder is found, however, common associations are with autoimmune disease, malignancy, dermatological disorders, pregnancy, and drugs.1,2

Two women aged 70 and 67 presented with a history of excessive bruising and soft tissue bleeding 2-5 months after starting clopidogrel (Plavix; Bristol-Myers Squibb, Sanofi-Synthelabo) for peripheral vascular disease. Their drugs had not changed recently in any other way. They had no clinical symptoms or signs of malignancy, antiphospholipid syndrome, or collagen vascular disease.

One patient had had a documented normal activated partial thromboplastin time at the time of starting clopidogrel; the other had not been tested. Investigation showed that the women had a normal platelet count, peripheral blood film, and prothrombin time. Both had a prolonged activated partial thromboplastin time of 48.6 and 77.6 seconds (normal range 23-33 seconds). Tests for lupus anticoagulant, anticardiolipin antibody, antinuclear factor, double-stranded DNA, and rheumatoid factor were negative. The women had low factor VIII (3.9 and 1.0 Bethesda units) within eight weeks of treatment. We treated both patients with 1 mg/kg of prednisolone. Concentrations of factor VIII rose to 119 and 136 IU/dl, and the inhibitor became undetectable (<0.4 Bethesda units) within eight weeks of treatment. The factor VIII inhibitor relapsed in one patient when the steroid dose was reduced, but we induced remission with azathioprine.

Corrections and clarifications

Variations and increase in use of statins across Europe: data from administrative databases

The authors of this paper, Tom Walley and colleagues, have alerted us to some errors in the data in their table (BMJ 2004;328:385-6). For Ireland, the total use of cerivastatin is in fact 0.416 (not 4.16) million defined daily doses (DDDs); this change affects the overall use of statins in Ireland, which becomes 11.06 (not 14.80) million DDDs a year. The value for DDDs per 1000 head of population in Italy is correct as stated. For France, the total use of atorvastatin is 256.58 (not 357.52) million DDDs, and the value for DDDs per head of population is 15.58 (not 23.56); these changes affect France's overall use of statins and its overall rate of use—these values become 730.46 (not 846.88) million DDDs and 48.11 (55.82) DDDs per head of population respectively.

Administering, analysing, and reporting your questionnaire

Readers may have been confused by the page reference given for reference 2 in this article by Tom Walley and colleagues, which became 730.46 (not 846.88) million DDDs and 48.11 (55.82) DDDs per head of population respectively.