

University of York Department of Health Sciences

Measurement in Health and Disease

## Increasing Response Rates to Postal Questionnaires

David Torgerson

Director, York Trials Unit

(Written version by Martin Bland)

### Background

Many, if not most, trials use postal questionnaires to collect outcome data on participants. Non-response to postal questionnaires can be a serious problem. Methods to reduce non-response need to be used.

Poor response to questionnaires will produce the following problems.

- Reduces statistical power of the study as the 'effective' sample size of study is reduced.
- Can introduce bias if non-response is systematically different between groups.

For every person that does not respond your trial will lose power. 20% non-response rates are typical. This means that you usually have to have 20% more participants recruited to make up for this loss. But more worryingly is bias. For example, Roberts *et al* undertook an RCT of paying women (£5) to return a questionnaire of HRT vs. no payment. The payment group had a 9% lower ever use of HRT than the no payment group ( $p = 0.05$ ). Because this was an RCT we know both groups will have the same HRT use. The difference is due to the poorer response rate (non-HRT users less likely to respond unless given an incentive). The Roberts study is a GOOD example of someone doing an MSc project based on survey data who also 'sneaked' in some proper research, a randomised trial.

### Methods of reducing non-response

A Cochrane review (Edwards *et al.* 2002;) has reviewed all the randomised trials of interventions to increase response rates. The review has identified a number of different ways of increasing response rates to surveys.

Questionnaire length has an influence on response. Long questionnaires do decrease response rates. For example, a single page questionnaire will produce a response rate of 67% compared with 50% for a 3 page questionnaire. But response rate is probably not linear (i.e., doubling from 3 to 6 won't have as a detrimental effect as doubling from 1 to 2).

**Table 1. Results of a questionnaire experiment (Iglesias and Torgerson 2000)**

Questionnaire Type	Response Rate
Short (4 pages)	48.9%
Medium (5 pages)	48.7%
Long (7 pages)	40.5%

P = 0.04 comparison of short vs. long.

**Table 2. Results of a factorial trial of payment and lottery (Roberts *et al.* 2000)**

Incentive	Response to questionnaire	P value, incentive vs. no incentive
Payment	67.6%	P = 0.01
No payment	56.1%	
Lottery	58.6%	P = 0.18
No lottery	53.7%	

Iglesias and Torgerson (2000) provide an example of a trial of questionnaire length. To test the effect of adding quality of life measures to an outcome questionnaire designed mainly to collect fracture data we undertook an RCT. Women 70 years and over were randomised to receive a short, a medium, or a long questionnaire. The short questionnaire (4 pages) asked about socio-demographics and questions on falls and fractures. The medium questionnaire (5 pages) was like the short questionnaire with the EuroQol added. The long questionnaire (7 pages) had both of the above plus the SF12. Table 1 shows the results. The long questionnaire had a significantly lower response rate than the short.

Monetary incentives to complete questionnaire can be offered. Unsurprisingly these always work, although ethics committees often dislike them seeing their use as a form of 'coercion'. Incentive is not linear; \$15 will only give 2.5 increase in response vs. \$1. Lotteries and prize draws are less effective or may be ineffective.

Roberts *et al.* (2000) reported a factorial trial of payment and incentives, comparing payment with no payment and a lottery with no lottery. The results are shown in Table 2.

This trial shows that direct payment works. It suggests that lotteries probably have weaker effects.

**Figure 1. Standard layout of the SF12.**

**Your Feelings**

(Please Circle One Number on Each Line)					
- These questions are about how you feel and how things have been with you <b>during the past month</b> . For each question please give the one answer that comes closest to the way you have been feeling.  How much during <b>the last month</b> :	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a) Have you felt calm and peaceful?	1	2	3	4	5
b) Did you have a lot of energy?	1	2	3	4	5
c) Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5

Other response enhancements (number of studies) which have been studied include:

- Use of coloured ink (1);
- Use of recorded delivery (6);
- Use of stamps instead of business reply (14);
- Use of first class post (1)
- Pre-contact (28)
- Follow up (12)
- More interesting questionnaire (2) and user friendly(1) and Factual vs. attitudinal (1)
- General questions last (1);
- University sponsorship (13)

**Questionnaire Design**

The layout or design of a questionnaire can have effects on response rates. Figure 1 shows the standard layout of the SF12. There are problems with this layout. Cynthia Iglesias noted that patients in a study of venous ulcers often made mistakes in completing the ‘standard’ SF12. She decided to alter layout to try to improve completion rates. The new SF12 layout is shown in Figure 2. Does layout make a difference? To test the effectiveness of the two layouts, Iglesias *et al* (2001) undertook an RCT. 1500 women aged 70+ years were randomised to be sent the standard or the changed version.

The overall response rates to the two questionnaires were the same. Item non-completion rates were significantly different. Using the standard SF12 26.6% of responses had one or more missing items, compared with 8.5% of responses using the modified SF12 (difference = 18.1%, 95% CI 11.1% to 25.1%).

**Figure 2. New layout for the SF12.**

6. During the **past 4 weeks**, how often have you **accomplished less** than you would have liked in your work or any other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?

(Please circle **one number only**)

All of the time    Most of the time    Some of the time    A little of the time    None of the time  
                          1                                    2                                    3                                    4                                    5

7. During the **past 4 weeks**, how often have you done work or other activities **less carefully** than usual **as a result of any emotional problems** (such as feeling depressed or anxious)?

All of the time    Most of the time    Some of the time    A little of the time    None of the time  
                          1                                    2                                    3                                    4                                    5

**Table 3. Some item response rates using two layouts for the SF12 (Iglesias *et al* 2001)**

SF12 question	Response rate to question		
	Standard	New	Difference, New minus Standard
Q1	99.1%	99%	-0.1
<b>Q2</b>	<b>96.8%</b>	<b>99%</b>	2.2
<b>Q3</b>	<b>93.2%</b>	<b>99%</b>	<b>5.8</b>
<b>Q4</b>	<b>94.6%</b>	<b>99%</b>	<b>4.4</b>
<b>Q5</b>	<b>86.4%</b>	<b>98.5%</b>	<b>12.1</b>
<b>Q6</b>	<b>94.6%</b>	<b>97.5%</b>	2.9
<b>Q7</b>	<b>89.6%</b>	<b>97.5%</b>	<b>7.9</b>
Q8	94.6%	98.5%	3.9

**NB – Those items in bold are the ones that were changed the others were already in a free standing format.**

Another question investigated was whether using single-sided printing of questionnaires produced a different response to double-sided printing. Anecdotally, we noted that some older respondents to double sided questionnaires ‘missed’ questions on the back. Puffer *et al* (2004) in a factorial trial tested whether or not single vs. double sided printing made a difference. They also tested whether a single or multiple booklet was better. In this trial, 3869 women were randomised to receive a single or a double sided questionnaire. This included SF12, EuroQol, and questions on medications and fractures. They were also randomised to receive a single vs. a multiple booklet. The study had more than 85% power to detect an absolute difference of 5%. The results were that the response to the single-sided questionnaire was 50.0%, compared to 47.4% for double-sided (95% CI of difference = -0.56% to 5.76%, P = 0.11). There were no differences in the number of

completed questions. For the single booklet the response rate was 50.3%, compared to 52.2% for the multiple booklet (95% CI of difference = -1.6% to 4.8%, P = 0.33). For the single booklet an average of 61 out of 68 items were completed, which was significantly greater than the average of 59 of the 68 items for the multiple booklets (P = 0.02).

These trials illustrate the importance of testing anecdotes. Cynthia's anecdotal experience was proven in a trial but Suezann's anecdotes were not. RCTs of questionnaire design are easy to do, particularly in the pilot phase of trials.

For another example, consider a trial of offering study results to participants. As part of her MSc project, a student sensibly undertook an RCT, comparing the response rates of trial participants who were offered the study results with those who were not offered (Cockayne). As part of the Calcium and Vitamin D trial 1,000 women as part of their final follow-up were randomised 3:1 to be offered the results on study completion or not to be told they would get their results. The aim was to see if this increased response rates for the final questionnaire. The results were that 94.3% responded when they were not offered the option of getting their results compared with 93.6% who were offered the results (no significant difference). 90% of those asked wanted a copy of the results.

Another thing which could possibly influence response is the position in the questionnaire of different types of question. Garratt carried out an RCT to compare quality of response for different positions of questions in the questionnaire within a back pain study. The SF36 and Roland and Morris back pain scales were randomly put either at the front or the back of a long questionnaire. Item response and internal consistency were measured. The SF36 had a mean of 0.56 missing items when placed at the back compared with 0.68 when placed at the front. Internal reliability was better for SF36 when placed at the front. The RDQ was not affected.

There is a suggestion that response rates tend to be a little lower if sensitive questions are used. One study found a reduction in asking about housing tenure (Windsor 1992) but not about ethnic origin, and another found no difference asking about sexual health (Barker and Cooper 1996)

What can we conclude from these studies? From the trial by Puffer we can conclude that it is best to use double sided and a single booklet. From Iglesias et al, we also need to change the SF12 layout to make it clearer. From Cockayne's study we know that offering results has no effect (should do so as good practice). Garratt study suggests best results are obtained if the QoL instrument is placed near the front of the questionnaire.

There are a few other principles which should help achieve good response rate. We should try to keep questionnaires short (easier said than done). We should put the crucial outcome measure up close to the front of questionnaire. For the main follow-up, sometimes I drop all secondary measures and just put in the main one to keep follow-up short. Reminders, keep it short. We could also consider an incentives for the final follow-up, such as £5 per questionnaire, not conditional on return. Telephone follow-up and telephone completion may be necessary to get information from non-responders.

## Discussion

There are a number of methods of improving response rates. High response rates are important to prevent bias and loss of power in trials. Before you use a questionnaire in a trial need to read systematic reviews of how to maximise your response rates and/or test them in a RCT.

Finally If you are not lucky enough to be doing a trial for your thesis (but a survey) you could still include an RCT of different methods of doing the survey.

## References

- Barker PJ, Cooper RF. Do sexual health questions alter the public's response to lifestyle questionnaires? *Journal of Epidemiology and Community Health* 1996; **50**: 688-688.
- Edwards P. Roberts I. Clarke M. DiGiuseppe C. Pratap S. Wentz R. Kwan I. Increasing response rates to postal questionnaires: systematic review. *British Medical Journal*. 2002; **324**: 1183.
- Iglesias CP, Birks YF, Torgerson DJ. Improving the measurement of quality of life in older people: the York SF-12. *QJM-an international journal of medicine* 2001; **94**: 695-698.
- Iglesias C. Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. *Journal of Health Services & Research Policy* 2000; **5**: 219-21.
- Puffer S. Porthouse J. Birks Y. Morton V. Torgerson D. Increasing response rates to postal questionnaires: a randomised trial of variations in design. *Journal of Health Services & Research Policy* 2004; **9**: 213-7.
- Roberts PJ. Roberts C. Sibbald B. Torgerson DJ. The effect of a direct payment or a lottery on questionnaire response rates: a randomised controlled trial. *Journal of Epidemiology & Community Health* 2000; **54**: 71-2.
- Windsor J. What can you ask about - the effect on response to a postal screen of asking about 2 potentially sensitive questions. *Journal of Epidemiology and Community Health* 1992; **46**: 83-85.