

## Increasing Response Rates to Postal Questionnaires

### 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.

### Non-response

- 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.

### Loss of power

- 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.

### Example of Bias

- Roberts et al undertook an RCT of paying women to (£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).

Roberts et al. 2000;54:71-72.

### The Roberts Study

- This 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 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.

Edwards et al. BMJ 2002;324:1183

## Questionnaire length

- Long questionnaires do decrease response rates.
- For example, a single page q'naire will produce a response rate of 67% compared with 50% for a 3 page q'naire.
- 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).

## Example of 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, medium or long questionnaire.

Iglesias & Torgerson. JHSRP 2000;5:219-21

## Questionnaire experiment

Questionnaire Type	Characteristic
Short (4 pages)	Socio-demographics + questions on falls and fractures
Medium (5 pages)	As above but also the EuroQol
Long (7 pages)	Both above plus SF12.

## Questionnaire Results

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.

## Monetary incentives

- Unsurprisingly these always work. Although ethics committees often dislike them seeing their use as a form of 'coercion'.
- Incentive not linear \$15 will only give 2.5 increase in response vs \$1
- Lotteries, prize draws are less effective or may be ineffective.

## Factorial trial of payment and incentives

Payment	67.6%	P = 0.01
No payment	56.1%	
Lottery	58.6%	P = 0.18
No lottery	53.7%	

Roberts et al. J Epid Comm Hlth 2000;54:71-2

## Incentives

- Direct payment works.
- Lotteries probably have weaker effects.

## Other response enhancements

- Use of coloured ink (1);
- Use of recorded delivery (6);
- Use of stamps instead of business reply (14);
- Use of first class post (1)
- Precontact (28)
- Follow up (12)
- More interesting q'naire (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.

## SF12 layout

### 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

## Problems with layout

- Cynthia Iglesias noted that patients in a study of venous ulcers often made mistakes in completing the 'standard' SF12.
- Decided to alter layout to try and improve completion rates.

## New SF12 layout

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)?

## Layout does it make a difference?

To test the effectiveness of the two layouts Iglesias et al undertook an RCT.

- 1500 women aged 70+ years were randomised to be sent the standard or the changed version.

## Results

- Overall response rates were the same.
- Item non completion rates were significantly different. The standard SF12 26.6% of responses had 1 or more missing items compared with 8.5% of the modified SF12 (difference = 18.1%, 95% CI 11.1% to 25.1%)

Iglesias et al. QJM 2001;94:695-98.

## Item response rates

	Standard	New	Diff
Q1	99.1%	99%	-0.1
Q2	<b>96.8%</b>	<b>99%</b>	2.2
Q3	<b>93.2%</b>	<b>99%</b>	<b>5.8</b>
Q4	<b>94.6%</b>	<b>99%</b>	<b>4.4</b>
Q5	<b>86.4%</b>	<b>98.5%</b>	<b>12.1</b>
Q6	<b>94.6%</b>	<b>97.5%</b>	2.9%
Q7	<b>89.6%</b>	<b>97.5%</b>	<b>7.9%</b>
Q8	94.6%	98.5%	3.9%

## Single vs Double sided

- Anecdotally we noted that some older respondents to double sided questionnaires 'missed' questions on the back.
- Puffer et al in a factorial trial tested whether or not single vs double sided printing made a difference. Also tested whether a single or multiple booklet was best.

Puffer et al. JHSRP; 2004: 213-17.

## Method

- 3869 women were randomised to receive a single vs double sided questionnaire (includes: SF12; EuroQol; questions on medications and fractures). Also single vs multiple booklet.
- Study had more than 85% power to detect an absolute difference of 5%.

## Results

Single Sided	50.0%
Double Sided	47.4%
95% CI of difference =	-0.56% to 5.76%, p = 0.11

There were no differences in number of completed questions

## Results

Multiple booklet	52.2%
Single booklet	50.3%
95% CI of difference =	-1.6% to 4.8%, p = 0.33

Single booklet average of 61/68 vs 59/68 items p = 0.02

## Importance of testing anecdote!

- Cynthia's anecdotal experience was proven in a trial but Suezann's anecdotes were not.
- RCTs of questionnaire design easy to do particularly in pilot phase of trials.

## Trial of offering study results

- 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.

## Method

- As part of the Calcium and 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.

## Results

- 94.3% responded when they were not offered the option of getting their results compared with 93.6% (no significant difference).
- 90% of those asked wanted a copy of the results.

## Questionnaire position

- In a RCT within a backpain study Garratt compared quality of response with position of questions in questionnaire. SF36 and Roland and Morris back pain were randomly put either at the front or the back of a long questionnaire. Item response and internal consistency were measured.

## Results.

- 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.

## Sensitive questions

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

Windsor, 1992, J Epi Comm Hlth 46:83-85

Barker & Cooper J Epi Com Hlth 1996;50:688

## Layout

- 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 QoL instrument is placed near the front.

## What do I do?

- Try to keep q'naire short (easier said than done).
- For the crucial outcome measure, put up close to front of questionnaire.
- For main follow-up sometimes I drop all secondary measures and just put in the main one to keep follow-up short.
- Reminders, keep short.
- Incentives for final follow-up (e.g., £5 per questionnaire not conditional on return).
- Telephone follow-up and telephone completion if necessary.

## 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 q'naire 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.