

# Mail surveys of general practice physicians: response rates and non-response bias

*James Young*

Statistician, Statistical Advisor for the Swiss Medical Weekly, Basel, Switzerland

In this edition of the Swiss Medical Weekly, Bergk and colleagues [1] consider whether mail surveys are still a useful tool in general practice research. Response rates from mail surveys can be so abysmally low that the results from such surveys are just not credible. They argue that the impact of reminders on survey results should be assessed early. If no bias can be detected, then one further reminder with a copy of the questionnaire should be sufficient in a well-run survey.

At first glance, this seems perfectly reasonable. After all, if one reminder does not change overall estimates of survey results, why continue with the process? The problem is that efforts to improve response rates in mail surveys, while necessary, are not sufficient for reliable estimates. Initial response rates in mail surveys are so low that efforts to improve response rates are unlikely to lead ultimately to a statistically adequate sample. In general it will also be necessary to adjust estimates for non-response bias. I will describe four approaches to adjusting estimates for non-response, concentrating on principles rather than giving formulae. Further details can be found in the references.

But first consider the survey presented by Bergk and colleagues. The initial response rate was 33%; their subsequent efforts nearly doubled this rate to a commendable 61%. While this is very good for a mail survey, a 60% response rate may still not be adequate from a statistical perspective. We have to expect that those that do not respond are different from those that do. 40% of respondents have not responded, and this example is a well-run survey. Often the situation will be much worse.

In table 1 there is good evidence that physicians who view drug interactions as a safety risk are more likely to respond. Table 1 also suggests that those physicians who are more aware of the issue – those that mention at least one clinically relevant interaction – are also more likely to respond. Would it be so surprising if those who think the issue important or with personal experience of the issue or with recent training were more likely to respond? Given that responses to other questions are likely to be correlated with these basic characteristics of respondents – their level of interest and awareness – are we then so sure we have reasonably unbiased estimates from a survey?

That is why it is necessary to adjust estimates for non-response. It is very difficult to know what sort of response rate is adequate. When I worked in official government statistics, as a rough rule of thumb we were reluctant to publish any data collected with a response rate below 70%. But an adequate response rate really depends on how different those who do not respond are from those who do – how different that is, in terms of their answers to survey questions. Ideally one would calculate a number of adjusted estimates for key questions under different, but quite plausible, non-response scenarios. If estimates appear relatively robust to a range of plausible non-response scenarios, then one can have confidence that non-response bias is not a problem.

The first approach to adjusting for non-response is to use time of return information [2–5]. For example, given the information in table 1, I would expect the percentage of physicians who regard drug interactions as a safety risk to be no more than 85% among the 784 physicians who never responded. In this case, my overall point estimate for all 2000 physicians will be at most 87% – that is  $(660 \times 0.91 + 136 \times 0.87 + 420 \times 0.85 + 784 \times 0.85) / 2000$ . Compare this with the 95% confidence intervals given in table 1: 89–93% for first respondents and 87–90% for all respondents. Whether the difference is important depends on context, but my point estimate is at the lower boundary of the interval calculated from all respondents. And my point estimate is probably too high because it is likely that the true percentage among those who never responded will be somewhat lower than 85%. More sophisticated use of time of return information might involve modelling the response as a (linear or curved) function of the order in which responses were received; predicting responses for the 784 non-respondents; and then estimating a proportion from all 2000 responses, both actual and predicted.

A second approach to adjusting for non-response is post-stratification [6]. Whether this approach will work or not depends on the information available about physicians in the population from which the sample was drawn. We need a variable that is known for all physicians and that is correlated both with responses and with the propensity to respond. This is not as difficult as it sounds.

For example, suppose we know the age of each physician; or better still, the number of years since graduation. Now suppose that younger physicians (or recent graduates) are more aware of the problems caused by drug interactions and are more likely to respond. Then age (or years since graduation) can be used to group respondents into a number of 'post-strata'. These groups should not be too small – say at least 30 or 40 respondents per group. We then compare the fraction of the respondents in each group with the fraction of the population in each group. A post-stratified estimate is a weighted average: a group average is given more weight when the group contains a lower fraction of respondents than it should have, and less weight when it contains a higher fraction of respondents than it should have.

A third approach to adjusting for non-response is to take a simple random sample of non-respondents after initial responses have been received [7]. The overall estimate is then a weighted average of the estimate for the initial respondents and the estimate for the sample of non-respondents. Only a small sample of non-respondents is needed – say again 30 or 40, so that the estimate from this sample is stable. But for this approach to work, nearly all the sampled non-respondents must be return a response and this means that face to face interviews are needed rather than just mailing out another questionnaire.

A fourth approach is the key question approach [8]. After the initial responses are received, each non-respondent is telephoned and asked just one or two quick questions; the key questions of the survey. Perhaps in this example the key questions are the first two: whether the physician re-

gards drug interactions as a safety risk and whether they think this risk is more or less important than other risks in treatment. Successfully used, this approach gives a high response rate for key survey questions. Since other questions in a survey tend to be correlated with key questions, regression techniques can be used to adjust estimates for the other survey questions.

With mail surveys, low response rates are the rule; not the exception. In a well-run survey, researchers should use one or more of these adjustment strategies in addition to at least one follow-up with a copy of the questionnaire. A classic mistake is putting too many resources into a large sample size and too few into managing non-response. A simple random sample of 2000 physicians is usually far larger than is really necessary. A simple random sample of just 400 physicians is sufficient to estimate a proportion to within  $\pm 5\%$  with 95% confidence. Instead of increasing the sample size, resources should be dedicated to improving follow-up and to adjusting estimates [9, 10]. If researchers take a larger sample, but fail to adjust estimates for the inevitable non-response, then their estimates will be precise – but not accurate.

---

*Correspondence:*

*James Young*

*Basel Institute for Clinical Epidemiology*

*University Hospital Basel*

*Hebelstrasse 10*

*CH-4031 Basel*

*Switzerland*

*jyoung@ubbs.ch*

---

## References

- 1 Bergk V, Gasse C, Schnell R, Haefeli WE. Mail surveys: obsolescent model or valuable instrument in general practice research? *Swiss Med Wkly* 2005;135:189–91.
- 2 Filion FL. Estimating bias due to nonresponse in mail surveys. *Publ Opin Quart* 1976;39:482–92.
- 3 Alho JM. Adjusting for nonresponse bias using logistic regression. *Biometrika* 1990;77:617–24.
- 4 Drane JW, Richter D, Stoskopf C. Improved imputation of non-responses to mailback questionnaires. *Statistics Med* 1993; 12:283–8.
- 5 Brenner H. Alternative approaches for estimating prevalence in epidemiologic surveys with 2 waves of respondents. *Am J Epidemiol* 1995;142:1236–45.
- 6 Smith TMF. Post-stratification. *Statistician* 1991;40:315–23.
- 7 Cochran, W. G., *Sampling techniques*. 3rd ed. New York: Wiley; 1977. p. 370–74.
- 8 Bethlehem JG, Kersten HMP. On the treatment of nonresponse in sample surveys. *J Off Statistics* 1985;1:287–300.
- 9 Fuller CH. Weighting to adjust for survey nonresponse. *Publ Opin Quart* 1974;38:239–46.
- 10 Wayne I. Nonresponse, sample size, and the allocation of resources. *Publ Opin Quart* 1976;39:557–62.

## The many reasons why you should choose SMW to publish your research

### What Swiss Medical Weekly has to offer:

- SMW's impact factor has been steadily rising, to the current 1.537
- Open access to the publication via the Internet, therefore wide audience and impact
- Rapid listing in Medline
- LinkOut-button from PubMed with link to the full text website <http://www.smw.ch> (direct link from each SMW record in PubMed)
- No-nonsense submission – you submit a single copy of your manuscript by e-mail attachment
- Peer review based on a broad spectrum of international academic referees
- Assistance of our professional statistician for every article with statistical analyses
- Fast peer review, by e-mail exchange with the referees
- Prompt decisions based on weekly conferences of the Editorial Board
- Prompt notification on the status of your manuscript by e-mail
- Professional English copy editing
- No page charges and attractive colour offprints at no extra cost

### Editorial Board

Prof. Jean-Michel Dayer, Geneva  
 Prof. Peter Gehr, Berne  
 Prof. André P. Perruchoud, Basel  
 Prof. Andreas Schaffner, Zurich  
 (Editor in chief)  
 Prof. Werner Straub, Berne  
 Prof. Ludwig von Segesser, Lausanne

### International Advisory Committee

Prof. K. E. Juhani Airaksinen, Turku, Finland  
 Prof. Anthony Bayes de Luna, Barcelona, Spain  
 Prof. Hubert E. Blum, Freiburg, Germany  
 Prof. Walter E. Haefeli, Heidelberg, Germany  
 Prof. Nino Kuenzli, Los Angeles, USA  
 Prof. René Lutter, Amsterdam, The Netherlands  
 Prof. Claude Martin, Marseille, France  
 Prof. Josef Patsch, Innsbruck, Austria  
 Prof. Luigi Tavazzi, Pavia, Italy

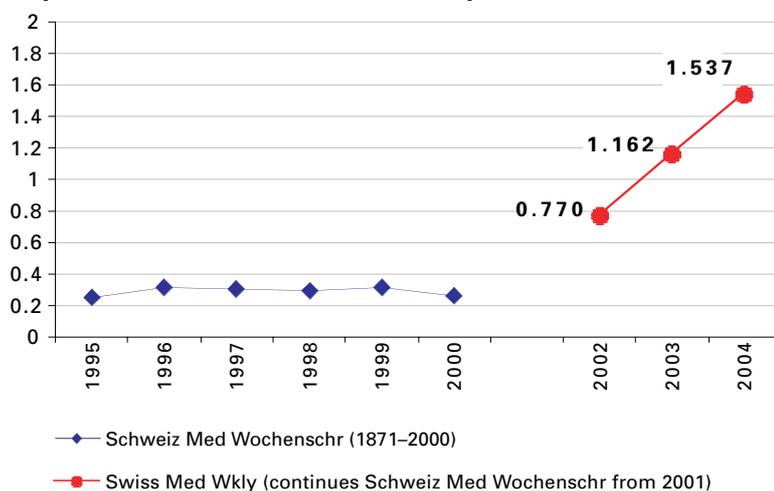
We evaluate manuscripts of broad clinical interest from all specialities, including experimental medicine and clinical investigation.

We look forward to receiving your paper!

Guidelines for authors:

[http://www.smw.ch/set\\_authors.html](http://www.smw.ch/set_authors.html)

### Impact factor Swiss Medical Weekly



All manuscripts should be sent in electronic form, to:

EMH Swiss Medical Publishers Ltd.  
 SMW Editorial Secretariat  
 Farnsburgerstrasse 8  
 CH-4132 Muttenz

Manuscripts: [submission@smw.ch](mailto:submission@smw.ch)  
 Letters to the editor: [letters@smw.ch](mailto:letters@smw.ch)  
 Editorial Board: [red@smw.ch](mailto:red@smw.ch)  
 Internet: <http://www.smw.ch>