

Dietary investigations – what are the effects of invalid selection procedures and measurement errors?

By Lars Berglund

Abstract

In observational studies where average levels or percentiles in a population distribution of nutritional components are to be assessed it is essential that the participating subjects are representative of the population to which the results are to be generalised. The optimum way of achieving this is by random selection from the target population in which all members have equal probability (or at least a known probability) of being chosen and also that all selected subjects participate in the study. The effects of selection bias and non-response can be large and are in general difficult to estimate. In the presentation of results it should always be stated how selection bias and non-response could have influenced the results.

In this paper a general selection procedure and certain problems with it is presented. A method for correction of nonresponse bias is also presented. The effect of measurement errors on correlation coefficients between nutritional components or between a nutritional component and other variables is discussed.

Introduction

In discussions of methods for population-based dietary assessment studies there is in general great concern about the validity of different assessment methods for measurement of the individual subject's true level. Different methods vary according to precision and bias. These considerations are important for their implications on bias and precision of estimates of average levels, percentiles and correlation coefficients. Less attention has been focused on the effects of how subjects have been selected from the population to which the results will be generalised. Substantial errors can result from biased selection procedures and low response rates on average levels, percentiles and correlation coefficients. Comparisons between different groups of subjects or between geo-

graphical areas or comparisons over time can be invalidated by differences in response rates or by low response rates where the reasons for non-response differ between studies.

One aim of this paper is to discuss selection bias and non-response bias in studies where the purpose is to give population-based estimates of average levels, percentiles or correlation coefficients for nutritional components. Another aim is to present a method for correction of non-response bias. A third aim is to discuss effects of measurement errors on correlation coefficients.

Concepts

The situation we are studying is observational studies such as cross-sectional or longitudinal studies and not intervention studies with random allocation to treatment groups. Some concepts will be used throughout the text and are defined here:

The target population is the collection of subjects to which we would like to generalise our results. The *frame population* is the collection of subjects that we can reach (e.g. persons in a register).

The under-coverage is the collection of subjects that belong to the target population but not the frame population and the *over-coverage* is the collection of subjects that belong to the frame population but not the target population. The sample is the subset of the frame population that is chosen for measurements.

Responders are the subjects in the sample on whom we have actual measurements after the study (Figure 1, see next page).

An observational study is a cross-sectional or longitudinal study where variable values are observed and not influenced.

Measurement error is defined as the difference between a subject's true value and the observed value. This means that the measurement error consists of the within-subject variation over time and the pure dietary assessment method error which can contain both a systematic and a random part.

The systematic error (or bias) is the lack of validity and the random error is the inverse of the reliability (precision) of the method.

High accuracy is defined as small bias and high precision.

Coverage errors

The problem of selection bias due to over-coverage is easy to overcome by not including in the sample subjects who when contacted are shown not to belong to the target population. This will lead to a smaller sample than was intended, but if the size of the over-coverage can be foreseen a larger sample can be taken to compensate for the subjects that are rejected. A special case of this situation is when a large screening is done to find dietary habits of subjects with a rare disease. In such a case the over-coverage can be over 99% of the frame population.

The under-coverage is more difficult to handle. Under-coverage can e.g. arise from a register that is not updated to contain the current target population. All statistical inference from the observed data will be made to the frame population (possibly after excluding the over-coverage as mentioned above). The validity of genera-

Lars Berglund, Statistician, BSc, Unit for Clinical Nutrition Research, Department of Public Health and Caring Sciences/Geriatrics, University of Uppsala, PO Box 609, SE-751 25 Uppsala, Sweden.
E-mail: Lars.Berglund@geriatrik.uu.se

This article is based on a lecture held at the conference "Dietary assessments – how can we interpret the results?" The conference was arranged by Uppsala Food and Nutrition Centre (*Uppsala Livsmedelscentrum*) in Uppsala on Nov 11, 1997.



lisations to other populations (target populations) must depend on other sources of information than the data in the sample. Any supplementary information that can be gathered about differences between the frame population and the target population may be helpful.

Non-response

Low response rate

The problem of non-response mainly consists of two parts. A low response rate will increase standard errors for estimates by reducing the number of subjects. This can be compensated for by dividing the desired sample size by the expected response rate (e.g. if 100 subjects are to be measured and the expected response rate is 80% the number of subjects to be contacted is $100/0.80 = 125$).

Bias

Non-response can also introduce a bias of unknown size. If the percentage of subjects eating less than 6.3 MJ/day (1,500 kcal/day) is to be estimated and the estimate from the responders is 30% and the response rate is 90%, the true value among all selected subjects is in the interval 27–37%. The length of this interval is the same as the non-response rate and this is also true in general. Note that this is not a confidence interval referring to the population but an interval showing the possible size of the non-response bias.

Reasons for non-response

There are several reasons for non-response but the two most common reasons are that subjects are not available during the study period and that subjects refuse to participate. In both cases there is a concern about non-response bias. People who are not available can be working more or travelling more than the average in the population and for these reasons have different diet habits. Subjects who refuse to participate will be a heterogeneous group, but there will be a tendency that people who are more interested in their diet in general will be more willing to participate in dietary studies and therefore there will be a difference between refusals and participants.

In general there will be a tendency that the probability for non-response for a subject is related to the subject's variable values, which means that the stratum in the population that are potential non-responders will differ (sometimes substantially) from the responder stratum.

Planning a large study

In planning a large study it is advisable to take part of the budget for a pilot study.

This can serve several purposes. The pilot study gives an estimate of the response rate and this information can be included when calculating the sample size in the main study. If several call-backs are done in the pilot study the information on the response rate for each call-back serves as information for the decision on how many call-backs shall be made in the main study to achieve a reasonable response rate. Another purpose of the pilot study, apart from giving information about the non-response problem, is to use measures of variability from it to calculate the sample size in the main study. This can be done so that a confidence interval will have a pre-specified length. The pilot study can also be used to test the feasibility of the measurement instrument, to test the questionnaire and to give training for interviewers.

Measuring the size of bias

To measure the size of non-response bias and to correct that bias, several methods exist and are described in the literature on sample survey methods. One method described by *Hansen and Hurwitz* (1) is to take a random subsample of the non-response subjects and make a major effort (presumably at a higher cost) to get measurements from everyone in the subsample. If we assume that the cost of measuring subjects in the subsample is 10 times the cost of measuring subjects in the first sample and the response rate in the first sample is 80% an optimum subsample size would be 30% of the non-responders. This will increase the total cost by 40% compared to the expected cost if all subjects were measured in the first sample, but the method will remove the non-response bias if all subjects in the

subsample are measured. The estimate of the population mean will be a weighted mean from the first sample mean and the subsample mean where the weights are the response rate and the non-response rate respectively from the first sample.

If this kind of subsampling is not feasible due to limitations in cost and time resources, there are two more indirect methods to assess the size of non-response bias.

Sociodemographic information

The first method makes use of the fact that there will be some sociodemographic information available in most studies for all selected subjects before they are contacted, e.g. sex, age, geographical area and sometimes occupation will be available if a register is the frame population. This information can be used to compare response and non-response groups. Large differences regarding sociodemographic variables can indicate large non-response bias if there is a low response rate, and results should then be interpreted with caution.

Figure 1. General selection procedure.

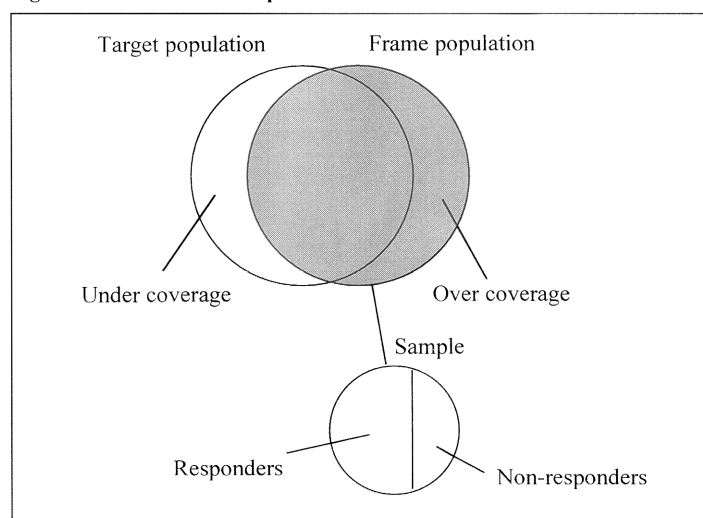


Table 1. Attenuation of correlation coefficients.

WBR ¹	Number of replicates				
	1	2	4	7	14
0.5	0.82	0.89	0.94	0.97	0.98
1.0	0.71	0.82	0.89	0.94	0.97
1.5	0.63	0.76	0.85	0.91	0.95
2.0	0.58	0.71	0.82	0.88	0.94
2.5	0.53	0.67	0.78	0.86	0.92

¹ Within-subject/between-subject variation ratio.

Differences in main variables

The second method makes use of differences in main variables between different call-backs. If there is a clear correlation (positive or negative) between main variables and the number of call-backs until the subject was measured there is a risk of a non-response bias and again interpretation of results should be done with care.

Papers reporting results from population-based dietary investigations do not always give information about response rates and reasons for non-response. In two studies using weighed food recording the response rates were 62% (2) and 53% (3), respectively, and in two other studies where food frequency questionnaires were used the response rates were 75% (4) and 69% (5), respectively. Only in one (2) of these four studies are the reasons for non-responses reported: 11% (29% of the non-response group) of the subjects were unavailable during the study period, 22% (59%) refused to participate and 4% (12%) had unreliable recordings.

Measurement errors

Random measurement error (within-subject variation and/or method error) will have implications for estimates of average levels and percentiles by increasing their standard errors, and the length of confidence intervals but will not introduce biases in those estimates.

Many dietary assessment methods also contain systematic errors when estimating

a subject's true level (e.g. underreporting of the total energy intake) and those errors will be transferred to estimates for the population.

Correlation coefficients

Another effect of random measurement errors in dietary studies is an attenuation of estimates of correlation coefficients between nutrient components and other variables or between different nutrient components. The size of this attenuation and methods to correct for it have been investigated by *Rosner* and *Willett* (6).

Table 1 shows the relative underestimation of correlation coefficients for combinations of within-subject/between-subject variation ratios (WBR), and the number of replicates for each subject. When the within-subject variation is twice the size of the between-subject variation, and there are two replicates per subject, an observed correlation coefficient will contain a systematic error in the size of 29% underestimation of the true correlation coefficient. This table assumes that there is random measurement error in only one of the two variables. If both variables are measured with error the attenuation effect will be even greater.

Table 1 can serve as a guide to judgement of the number of replicates needed for minimising the attenuation effect of random measurement errors. The input needed (the WBR) can be calculated from a pilot study if some of the subjects or all of them are measured at least twice.

Conclusions

More effort should be made in reports of results from dietary investigations to define the population to which results will be generalised. It should also be stated whether the main purpose is to estimate average levels and percentiles or correlation coefficients. In the first case the main concern would be non-response biases and their size and what measures have been taken to overcome the problem. In the latter case it should be described to what extent the within-subject variation has been dampened by taking replicate measurement for each subject.

Acknowledgements

The author would like to thank *Rawya Mohsen* and *Annika Smedman* for comments on an earlier draft.

References

1. Hansen MH, Hurwitz WN: The problem of non-response in sample surveys. *J Am Stat Assoc* 1946;41:517-29.
2. Shortt CT, Duthie GG, Robertson JD, Morrice PC, Nicol F, Arthur JR: Selenium status of a group of Scottish adults. *Eur J Clin Nutr* 1997; 51:400-4.
3. Harbottle L, Duggan MB: Dietary assessment in Asian children – a comparison of the weighed inventory and diet history methods. *Eur J Clin Nutr* 1993;47:666-72.
4. Jeffery RW, French SA, Forster JL, Spry VM: Socio-economic status differences in health behaviors related to obesity: the Health Worker Project. *Int J Obesity* 1991;15:689-96.
5. Bolton-Smith C, Woodward M: Antioxidant vitamin adequacy in relation to consumption of sugars. *Eur J Clin Nutr* 1995;49:124-33.
6. Rosner B, Willett WC: Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol* 1988; 127:377-86. □