

# How much loss to follow-up is acceptable in long-term randomised trials and prospective studies?

Mary S Fewtrell,<sup>1</sup> Kathy Kennedy,<sup>1</sup> Atul Singhal,<sup>1</sup> Richard M Martin,<sup>2</sup> Andy Ness,<sup>3</sup> Mijna Hadders-Algra,<sup>4</sup> Berthold Koletzko,<sup>5</sup> Alan Lucas<sup>1</sup>

There is increasing evidence that health and development in adult life is influenced or “programmed” by factors, including nutrition, operating during foetal life and infancy. Early nutrition in a variety of animal species, including primates, has been demonstrated to influence later outcomes, including the likelihood of cardiovascular disease, learning and behaviour problems and longevity (see supplementary data).

Evidence of programming in humans has, until recently, come largely from historical observational studies that showed associations between small size in early life and adult disease risk (see supplementary data). These cohorts were constructed from available health records. They by necessity relied on indirect measures of maternal and infant nutrition and often lacked detailed data on potential confounding variables. Many cohorts enrolled people who were born in the first half of the 20<sup>th</sup> century, and it is possible that the nature and size of any associations are different in contemporary populations. Thus, although these studies have generated considerable interest they have been unable to examine direct associations with diet or establish whether associations are causal and cannot, therefore, be used to inform infant feeding recommendations. This has led to a greater emphasis on the need for RCTs

to test early nutritional interventions and prospective observational cohorts.

RCTs are generally accepted as methodologically the best approach for informing health policy. They can equalise unknown as well as known confounding factors and so can demonstrate causation; they permit estimation of effect size and so can be used to assess likely economic benefits; and they can, if adequately powered, measure expected adverse effects and thus address safety. Nevertheless, in the context of nutritional programming of disease later in life, they have certain limitations. For example, some trials cannot be performed because they would be unethical (eg, breast feeding versus formula feeding). In addition, although certain measures during childhood or adolescence are predictive of final adult outcome (eg, cognitive function), other diseases will not become apparent for decades, necessitating the use of “proxies” of later disease risk that can be measured at younger ages.

Contemporary prospective observational studies are also recognised to be important in investigating nutritional programming in humans and are complementary to RCTs. They identify defined (often large) populations, measure them precisely and follow them up longitudinally. If the studies are population based they might also be more generalisable. They have better measures of exposure and confounding factors than historical cohort studies, and, of particular relevance to programming research, they may include detailed early physiological, biological and social data, allowing more complete adjustment for confounding factors.

Given the obvious requirement for long-term follow-up in the investigation of nutritional programming of health outcomes, an important issue has become

increasingly apparent, affecting both RCTs and prospective observational studies, namely, cohort attrition or loss to follow-up. This is an important issue for the field, with implications for study design and funding. The aim of this paper is to consider the statistical implications of attrition in both RCTs and cohort studies to identify the most effective ways of dealing with attrition when analysing and reporting studies. Factors influencing follow-up rates, measures that can be taken to minimise attrition in future studies and potential alternative approaches to the problem are discussed in the appendices. The paper uses examples from studies examining the effects of nutritional interventions or early-life nutrition on later outcomes, but many of the conclusions are applicable to studies in other fields.

## WHAT IS AN “ACCEPTABLE” FOLLOW-UP RATE?

There are no universally agreed criteria for acceptable follow-up rates in nutritional RCTs or cohort studies. In RCTs, typically investigating drugs or other therapies, it has been suggested that a loss to follow-up  $\leq 5\%$  is usually of little concern, whereas a loss of  $\geq 20\%$  poses serious threats to validity, with in-between rates leading to intermediate levels of problems.<sup>1</sup> Indeed, a cut-off of 80% is used in Evidence-Based Medicine (EBM) “Levels of Evidence” to separate “high”- and “low”-quality randomised trials.<sup>2</sup> This figure is based on the concept of being able to detect a hypothesised difference between randomised groups at follow-up after applying the “worst case scenario” for missing data – that is, assuming that subjects lost to follow-up from each arm of the study have the outcome seen in the other limb. Although this approach can easily be applied to studies with dichotomous outcome measures (such as death or survival), its applicability to continuous physiological variables, such as blood pressure, fat mass or skin-fold thicknesses, commonly used as outcomes in nutritional programming studies is uncertain. Rates of 50–80% follow-up have been suggested as acceptable by different authors in the context of epidemiological cohorts, although in most cases the validity of these recommendations has not been tested.<sup>3,4</sup>

## WHAT INFLUENCES FOLLOW-UP RATES?

Our own RCTs of infant nutritional interventions have generally achieved

<sup>1</sup> Childhood Nutrition Research Centre, UCL Institute of Child Health, London WC1N 1EH; <sup>2</sup> Department of Social Medicine, University of Bristol; <sup>3</sup> Department of Oral and Dental Science and Department of Social Science, University of Bristol; <sup>4</sup> University Medical Centre, Groningen, The Netherlands; <sup>5</sup> University of Munich, Dr von Hauner Childrens’ Hospital, Germany

**Correspondence to:** Dr Mary Fewtrell, Reader in Childhood Nutrition, Childhood Nutrition Research Centre, UCL Institute of Child Health, London WC1N 1EH; [m.fewtrell@ich.ucl.ac.uk](mailto:m.fewtrell@ich.ucl.ac.uk)

high short-term follow-up rates, typically 80–90% at 18 months,<sup>5–8</sup> which is in the range that would be considered “acceptable” for the purposes of EBM. However, longer-term follow-up rates in our studies and those reported from other groups working on the programming of adult diseases clearly depend on a number of factors including the age of the subject, nature and perceived benefit of the test, degree of inconvenience involved and the ability to trace and contact subjects. These points are illustrated in figs 1A and 1B, which show follow-up rates attained in different cohort studies and RCTs according to the age at follow-up and the nature of the tests used. Further examples are provided in Appendix 1 in the supplementary data.

### WHAT ARE THE STATISTICAL CONSEQUENCES OF ATTRITION?

Although we acknowledge the importance of aiming for maximum follow-up in any study, in practice it is inevitable that losses to follow-up will occur, and these are likely to increase with time. Given this reality, and assuming that obtaining follow-up data from RCTs and cohorts in the field of nutritional programming is considered to be worthwhile, we suggest that, rather than setting a fixed level of what constitutes “acceptable” follow-up, it is more helpful to consider the statistical implications of loss to follow-up and ways in which these issues can be most clearly presented and discussed in publications.

Altman<sup>9</sup> suggested that “assessing whether a trial was a good one should take account of circumstances, including

what is achievable. In trials of lifestyle interventions, for example, such as dieting or smoking cessation, such drop-out rates (ie, <20%) are rarely achieved, unless using an unrealistically short follow-up period”. The latter consideration clearly applies to work in the field of nutritional programming. Altman further suggests that “while there is potential value in guidelines, these should not in general be interpreted as rules, and we should not disguise the fact that exercising judgement is a major element of statistics”.

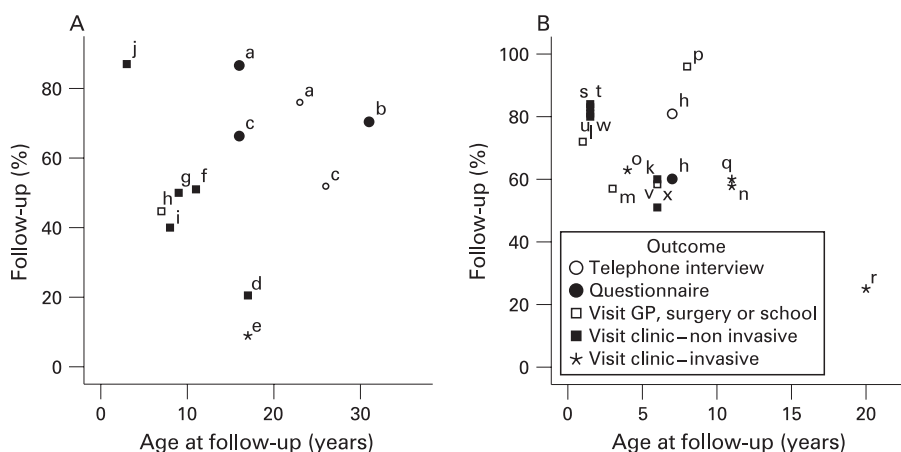
Attrition is important statistically for three principal reasons – its effect on study power, bias and generalisability:

1. Study power: Reduced sample size can affect the power of the study to detect a hypothesised difference. Indeed, it might be possible from the data presented in fig 1 to predict whether the likely sample size available for a follow-up study will be adequate to detect the hypothesised difference in outcome. If there is not a reasonable expectation of attaining the required sample size, it is questionable whether it is ethical to proceed with the study; although in some circumstances follow-up of an underpowered trial might be justified if data is to be pooled with those from similar studies using standardised protocols. We would argue that loss of sample size may not always present a serious problem for the physiological outcomes typically used in nutritional programming studies (eg, blood pressure or plasma lipid concentrations), where fairly large effect sizes are anticipated, requiring generally modest samples. For example, observed differences in cardiovascular risk factors between groups randomised to different

infant diets are typically in the range 0.5–0.7SD, requiring a sample size of approximately 64–100 per group to achieve 80% power at 5% significance.<sup>10–13</sup> Nevertheless, it is important that, where possible, trial sizes are large enough to exclude smaller effect sizes that might also have public health significance. It is also important to appreciate that, although studies should clearly be adequately powered at initiation, including an allowance for attrition, some outcomes now being examined at long-term follow-up in programming research are not necessarily those for which the study was initially powered, as the focus of scientific interest may have shifted with developments in the intervening years.

2. Bias can be defined as any systematic error in a study that results in an incorrect estimation of the association between exposure and outcome. Attrition introduces a form of selection bias, since loss to follow-up is rarely a truly random event. Dumville *et al*<sup>14</sup> considered the issue of reporting attrition in RCTs and argued that researchers should be more explicit about loss to follow-up and present tables of baseline data separately for those seen or not seen. If baseline characteristics are found to differ between those seen and not seen at follow-up, or if a potentially important predictor variable is more unbalanced between randomised groups at follow-up than at baseline, this may suggest bias. Such factors could be included as covariates in the analyses. However, although important, this may still be inadequate since imbalance of measured characteristics often implies imbalance of unmeasured characteristics.

Kristman *et al*<sup>5</sup> addressed the statistical effects of attrition in cohort studies using a simulation study. They found no important bias even with losses of up to 60% when data were “missing completely at random” or “missing at random”. The “missing at random” category assumed that dropouts were related to a characteristic measured at baseline or at subsequent follow-up (for example, socioeconomic status) but not to the outcome variable being measured. Since relevant baseline characteristics can be incorporated as covariates in analyses, this type of missing data would be considered to be “ignorable”. By contrast, the simulation suggested that when data were “not missing at random” (that is, dropouts were related to unobserved information or to the outcome variable), even small losses to follow-up (as little as 20%) could result in considerable bias in the results. In practice, it is, of course, impossible to identify when



## Leading article

loss to follow-up is related to unmeasured variables.

3. Generalisability can be defined as the extent to which research findings can be applied to settings other than the study sample in which they were tested. It is important to recognise that generalisability is an issue in many trials with excellent follow-up rates and a low risk of bias. For example, trials in the field of cardiovascular disease are often conducted in males, excluding the extremes of age, and their generalisability to the whole population and to real-life clinical settings might be questioned.

In practical terms, attrition may affect the generalisability of the results to the wider population either by introducing bias, which affects the ability to draw the correct conclusion in the study population and hence the general population. Even in the absence of bias, however, attrition can result in loss of power to a degree that affects the ability to draw a robust conclusion in the study population, which in turn affects generalisability. The relevance of attrition to generalisability is arguably a greater problem in cohort studies than in RCTs since the former are generally designed to be representative of a particular population. The effect of non-representative population sampling in a RCT testing for large physiological effects in an already selected group of subjects is perhaps less of an issue. The decreasing initial participation rates seen in large cohort studies over the past two decades, combined with subsequent attrition may pose a further threat to generalisability, although non-participation resulted in minimal bias in estimates of relative risk in one large cohort.<sup>15</sup>

### Attrition and effect size

It is important to consider whether the observed effect size might itself be influenced by attrition. Using data from the limited number of studies in fig 1B, we found no evidence that standardised effect sizes were greater in studies with higher attrition rates. It should be noted that this analysis is itself potentially flawed since (i) it involves comparing standardised effect sizes for different outcomes (sometimes studies find a large effect size for one outcome combined with a small effect size for another); (ii) it is possible that the effect size for certain outcomes genuinely amplifies with time, so any apparent relationship between attrition rate and effect size may not be causal; (iii) it is possible that studies with high attrition rates and a large effect size are more likely to be published than

### Box 1 Summary points

- ▶ Long-term follow-up of randomised controlled trials (RCTs) and prospective observational cohorts is an essential component of research into nutritional programming in humans.
- ▶ Loss to follow-up is inevitable with time, even with the best study design and conduct.
- ▶ The statistical consequences of attrition – loss of power, bias and generalisability, may vary depending on the design and aims of the study.
- ▶ The issue of attrition should be explicitly discussed in the context of each study.
- ▶ It is unnecessary and unhelpful to use fixed levels for “acceptable” follow-up rates, and this practice may curtail future research in this field.

those with small effect sizes. This analysis is necessarily limited given the small number of publications reporting long-term follow-up of RCTs in the programming field. More comprehensive consideration, however, of the influence of study characteristics on treatment effects has been conducted and could be used as a basis for practice in future.<sup>16</sup>

It is also important to consider the potential for attrition to introduce spurious differences between randomised groups, if the degree of bias resulting from attrition differs between groups. This can be addressed by performing a sensitivity analysis. In its simplest terms this could estimate how different the result would have to be in subjects not seen at follow-up in order to negate the difference between groups observed in those who were successfully studied.

Another important consideration in RCTs is whether the intervention itself results in differential attrition and consequent bias between groups, which would affect the continued validity of the randomisation for later outcomes. This can be

easily examined and has not generally been observed in follow-ups of nutritional intervention trials to date.<sup>5-7 10 13 17-19</sup>

### CONCLUSIONS

Long-term follow-up of RCTs and observational cohorts is an essential component of research into nutritional programming in humans and in other areas of research. Loss to follow-up is inevitable with time, even with the best study design and conduct. The potential effects of attrition should be explicitly acknowledged and dealt with when follow-up studies are analysed and reported, in the context of the aims of the study. In particular, it is important to address the issues of bias and generalisability of findings as far as possible and, in the case of RCTs, to examine whether the intervention has influenced attrition. The adequacy of the sample size attained at follow-up to detect the hypothesised effect should also be addressed, particularly when there are negative findings and there is substantial attrition. In general,

### Box 2 Suggested minimum reporting requirements for addressing attrition in long-term follow-up studies

1. Provide clear, unambiguous information on the flow of subjects through the study/cohort at each stage. Give explicit information on the attrition rate and the attrition in each group for a randomised trial.
2. Discuss the ability of the follow-up study to detect the hypothesised outcome effect with the sample size attained.
3. Discuss the potential for attrition to have introduced bias. Do baseline and/or measured variables differ in those seen and not seen? Provide baseline characteristics for those seen and not seen at follow-up separately for each intervention group.
4. Discuss whether attrition is likely to have affected the generalisability of the findings to the original study population (which may or may not have been representative of the larger population) and to the general population.
5. Provide an appropriate sensitivity analysis; describe the assumptions on which it is based.\*

\*In its simplest terms this could estimate how different the result would have to be in subjects not seen at follow-up to negate the difference between groups observed in those who were successfully studied.

results should be presented as estimates of effect size with confidence limits. In all cases, the interpretation of results should include a discussion of the potential effects of attrition. Suggested reporting requirements relating to attrition are listed in Box 2: more general and comprehensive reporting requirements for observational studies have been developed by the STROBE group.<sup>20</sup> We believe that, providing these issues are acknowledged and addressed, it is unreasonable and unnecessary to use cut-offs for “acceptable” follow-up rates, such as those proposed for short-term pharmaceutical trials. Indeed, the application of such criteria will effectively curtail further research into the long-term effects of early nutrition on later health outcomes and potentially prevent the development of evidence-based guidelines for pregnant women and their offspring.

**Acknowledgements:** The authors are partners in the European Union FP-6 Early Nutrition Programming Project Consortium and receive funding to conduct follow-up studies of RCTs and observational cohorts (Food-CT-2005-007036). The authors would like to acknowledge helpful comments made by Tim Cole, Jorn Olsen and Peter Whincup.

**Competing interests:** None.

► Additional text and references are published online at <http://adc.bmj.com/content/vol93/issue6>

Accepted 19 November 2007

*Arch Dis Child* 2008;**93**:458–461.  
doi:10.1136/adc.2007.127316

## REFERENCES

1. **Sackett DL**, Strauss SE, Richardson WS, *et al*. *Evidence-based medicine: how to practice and teach EBM*. Edinburgh: Churchill Livingstone, 2000.
2. **Oxford CEBM**. [http://www.cebm.net/levels\\_of\\_evidence.asp](http://www.cebm.net/levels_of_evidence.asp) (accessed 26.2.07).
3. **Kristman V**, Manno M, Cote P. Loss to follow-up in cohort studies: how much is too much? *Eur J Epidemiol* 2004;**19**:751–60.
4. **Babbie ER**. *Survey research methods*. Belmont, CA: Wadsworth, 1973.
5. **Fewtrell MS**, Morley R, Abbott RA, *et al*. Double-blind randomised trial of long-chain polyunsaturated fatty acid supplementation in formula-fed to preterm infants. *Pediatr* 2002;**110**:73–82.
6. **Fewtrell MS**, Abbott RA, Kennedy K, *et al*. Randomized double-blind trial of LCPUFA-supplementation using fish oil and borage oil in preterm infants. *J Pediatr* 2004;**144**:471–79.
7. **Lucas A**, Stafford M, Morley R, *et al*. Efficacy and safety of long-chain polyunsaturated fatty acid supplementation of infant-formula milk: a randomised trial. *Lancet* 1999;**354**:1948–54.
8. **Fewtrell MS**, Morley R, Abbott RA, *et al*. Catch-up growth in small-for-gestational-age term infants: a randomized trial. *Am J Clin Nutr* 2001;**74**:516–23.
9. **Altman DG**. Statistics in medical journals: some recent trends. *Stat Med* 2000;**19**:3275–3289.
10. **Singhal A**, *et al*. Promotion of faster weight gain in infants born small for gestational age. Is there an adverse effect on later blood pressure? *Circulation* 2007;**115**:213–20.
11. **Singhal A**, Cole TJ, Lucas A. Early nutrition in preterm infants and later blood pressure: two cohorts after randomised trials. *Lancet* 2001;**357**:413–19.
12. **Singhal A**, Farooqi IS, O’Rahilly S, *et al*. Early nutrition and leptin concentrations in later life. *Am J Clin Nutr* 2002;**75**:993–9.
13. **Forsyth JS**, Willatts P, Agostoni C, *et al*. LCPUFA supplementation in infant formula and later BP in childhood: follow up of a randomized controlled trial. *BMJ* 2003;**326**:953–8.
14. **Dumville JC**, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. *BMJ* 2006;**332**:969–71.
15. **Nohr EA**, Frydenberg M, Henriksen TB, *et al*. Does low participation in cohort studies induce bias? *Epidemiology* 2006;**17**:413–8.
16. **Sterne JAC**, Juni P, Schulz KF, *et al*. Statistical methods for assessing the influence of study characteristics on treatment effects in ‘meta-epidemiological’ research. *Stat Med* 2002;**21**:1513–24.
17. **Singhal A**, *et al*. Infant nutrition and stereoacuity at ages 4–6yrs. *Am J Clin Nutr* 2007;**85**:152–9.
18. **Scott DT**, Janowsky JS, Carroll RE, *et al*. Formula supplementation with LCPUFAs: Are there developmental benefits? *Pediatr* 1998;**102**:e59.
19. **Kalliomaki M**, Salminen S, Poussa T, *et al*. Probiotics and prevention of atopic disease: 4yr follow-up of a randomised controlled trial. *Lancet* 2003;**361**:1869–71.
20. **von Elm E**, Altman DG, Egger M, *et al*. STROBE Initiative: Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;**335**:806–8.

## Archivist

### Long-term effect of repeated antenatal steroid

A single course of steroid given to women at risk of preterm delivery improves neonatal morbidity and mortality. If delivery is more than 7 days after steroid administration, however, perinatal mortality is increased and birth weight decreased. The effects of repeated courses of steroid are uncertain though there is evidence that they may have a beneficial effect on the neonatal lung but reduce birth weight. Now two follow-up studies of randomised trials in Australia and New Zealand and in the USA have been reported (Caroline A Crowther and colleagues. *New England Journal of Medicine* 2007;**357**:1179–89 and Ronald J Wapner and colleagues, *ibid*: 1190–8; see also editorial, *ibid*: 1248–50).

A total of 982 women (1146 fetuses) in Australia and New Zealand and 495 (594 fetuses) in the USA, all at risk of delivery before 32 weeks, were randomised to weekly betamethasone or placebo  $\geq 7$  days after first steroid administration. The weekly dose of betamethasone was higher in the USA (12 mg on two successive days vs 11.4 mg on one day only). The short-term beneficial results were reported in the *Lancet* (Australian study) and in the *American Journal of Obstetrics and Gynecology* (US study) in 2006.

Totals of 1047 and 556 children were assessed at 2–3 years of age. Neither study showed significant differences between the betamethasone or placebo groups in rates of major neurodevelopmental problems nor in body size or blood pressure. Attention problems were more common among the repeated steroid group in Australia and New Zealand and there was a statistically nonsignificant increase in cerebral palsy in the repeated steroid group (2.9% vs 0.5%,  $p = 0.12$ ) in the USA. These researchers come to different conclusions. Crowther and colleagues suggest that clinicians “may wish to consider” repeat steroid administration whereas Wapner and colleagues conclude that their findings (a possible increase in risk of cerebral palsy) “argue against” this policy. Could the higher dose of betamethasone in the American study have affected the results? Only further research will tell.



## How much loss to follow-up is acceptable in long-term randomised trials and prospective studies?

Mary S Fewtrell, Kathy Kennedy, Atul Singhal, et al.

*Arch Dis Child* 2008 93: 458-461  
doi: 10.1136/adc.2007.127316

---

Updated information and services can be found at:  
<http://adc.bmj.com/content/93/6/458.full.html>

*These include:*

### References

This article cites 17 articles, 9 of which can be accessed free at:  
<http://adc.bmj.com/content/93/6/458.full.html#ref-list-1>

Article cited in:  
<http://adc.bmj.com/content/93/6/458.full.html#related-urls>

### Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

### Topic Collections

Articles on similar topics can be found in the following collections

[Childhood nutrition](#) (403 articles)  
[Childhood nutrition \(paediatrics\)](#) (211 articles)  
[Infant nutrition \(including breastfeeding\)](#) (226 articles)  
[Child health](#) (2179 articles)  
[Infant health](#) (354 articles)  
[Pregnancy](#) (248 articles)  
[Reproductive medicine](#) (535 articles)  
[Epidemiologic studies](#) (1001 articles)  
[Diet](#) (206 articles)  
[Drugs: cardiovascular system](#) (287 articles)

---

### Notes

---

To request permissions go to:  
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:  
<http://group.bmj.com/subscribe/>