

Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders

Dieter Wolke, Andrea Waylen, Muthanna Samara, Colin Steer, Robert Goodman, Tamsin Ford and Koën Lamberts

Background

Participant drop-out occurs in all longitudinal studies, and if systematic, may lead to selection biases and erroneous conclusions being drawn from a study.

Aims

We investigated whether drop out in the Avon Longitudinal Study of Parents And Children (ALSPAC) was systematic or random, and if systematic, whether it had an impact on the prediction of disruptive behaviour disorders.

Method

Teacher reports of disruptive behaviour among currently participating, previously participating and never participating children aged 8 years in the ALSPAC longitudinal study were collected. Data on family factors were obtained in pregnancy. Simulations were conducted to explain the impact of selective drop-out on the strength of prediction.

Results

Drop out from the ALSPAC cohort was systematic and children who dropped out were more likely to suffer from disruptive behaviour disorder. Systematic participant drop-out according to the family variables, however, did not alter the association between family factors obtained in pregnancy and disruptive behaviour disorder at 8 years of age.

Conclusions

Cohort studies are prone to selective drop-out and are likely to underestimate the prevalence of psychiatric disorder. This empirical study and the simulations confirm that the validity of regression models is only marginally affected despite range restrictions after selective drop-out.

Declaration of interest

None.

Prospective studies provide one of the strongest methodologies for studying aetiological mechanisms,¹ but are vulnerable to selection biases as a result of losses to follow-up. Participant losses can be random² or systematically related to social or biological characteristics of the participants that may or may not be associated with the outcome of interest.^{3,4} If there is systematic loss to follow-up related to the potential aetiological factors under investigation, any conclusions drawn from the study may be erroneous. We investigated the impact of selective participant drop-out using a prospective study and conducted a series of simulations to explain the empirical findings. The Avon Longitudinal Study of Parents And Children (ALSPAC) collected data about disruptive behaviour problems from teachers on all children attending participating schools within the Avon area at 7 years 9 months allowing us to examine the following questions. First, do children continuously participating in the longitudinal cohort (current ALSPAC) differ from children going to the same schools who were never part of the cohort (never ALSPAC)? Second, do those who have dropped out of the cohort (previous ALSPAC) differ systematically from those who stayed on (current ALSPAC)? Third, are the prediction models for disruptive behaviour disorders the same for those who are currently still participating in the study (current ALSPAC) compared with those who dropped out (previous ALSPAC)? Finally, we conducted simulations to explain the impact of selective drop-out on the strength of prediction if drop-out, predictor and criterion variables are correlated to varying degrees.

Method

Participants

The Avon Longitudinal Study of Parents And Children⁵ is a population-based study which investigates a wide range of environmental, genetic and psychosocial influences on the health and development of children and their parents. Figure 1 illustrates

participation in ALSPAC up to and including the data gathered from teachers when the children were in school year 3. The 14 541 pregnant mothers recruited into the study between April 1991 and December 1992 had 14 062 live births. At 1 year 13 988 infants were alive and 13 971 at 7 years of age. When compared with 1991 national census data, the ALSPAC sample was found to be similar to the UK population as a whole, having only a slightly higher proportion of married or cohabiting mothers who were owner-occupiers and who had a car in the household. There were also a slightly smaller proportion of mothers from ethnic minority groups.⁵

At 7 years 9 months, as part of a study on disruptive behaviour disorders (attention-deficit hyperactivity disorder (ADHD) and behaviour disorders), teachers in the geographically defined study area (the old county of Avon in the UK) were asked to complete the Development and Well-Being Assessment (DAWBA)⁶ on all the children in their class with a birth date between April 1991 and December 1992. From a total of 10 431 children eligible to be contacted, teachers returned questionnaires for 3975 children whose parents also participated in this survey (current ALSPAC children), and 1140 children who had participated in previous parts of the ALSPAC study but whose parents did not respond to the current survey (previous ALSPAC children) (Fig. 1). The teacher completion was thus 5115/10 431 of eligible children (49%) or 5115/13 971 of all survivors (37%). In addition, teacher data was returned for 4383 children who had never been recruited into the ALSPAC study or had moved into the area after the study had started (never ALSPAC children). The study was approved by the ALSPAC Ethics and Law Committee and local research ethics committees.

Procedures

During pregnancy, and annually since then, detailed information about the mothers and their partners has been collected via

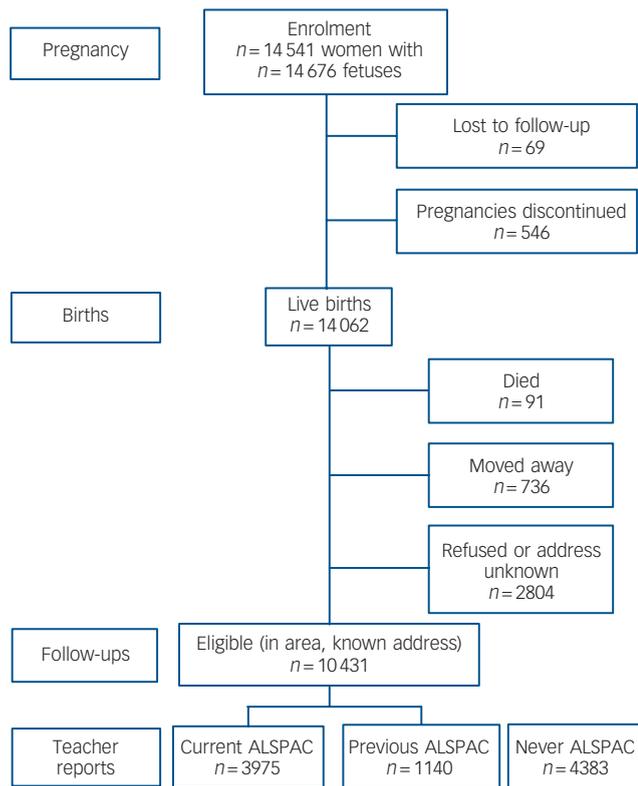


Fig. 1 Description of ALSPAC sample: flow chart.

self-report questionnaire with regard to medication, symptoms, diet and lifestyle, attitudes and behaviour, and social–environmental features.⁵ From 4 weeks after the birth of the child, mothers completed questionnaires about the child’s health, development and environment (biannually on average).

When the children were 7 years and 9 months, teachers were asked to complete the Development and Well-Being Assessment (DAWBA)⁶ as part of a study on disruptive behaviour disorders. The teacher version of the DAWBA is a brief structured questionnaire that covers the operationalised diagnostic criteria for the main disruptive behavioural disorders included in DSM–IV,⁷ namely oppositional defiant disorder, conduct disorder and ADHD. Thirty-nine cases were excluded where there were insufficient data from teachers for a diagnosis to be made.⁸

Data collection from teachers occurred over three academic years (1999, 2000 and 2001), with response rates varying from year to year. A minority of schools declined to participate (5%, 13% and 6% respectively) and some failed to respond to the invitation (17%, 37% and 16%), but the response rate from the schools who agreed to participate was high (80%, 99% and 80%) leading to an overall response rate of 62%, 50% and 63% for each year.

The following family-based risk factors were assessed during pregnancy: marital status (married *v.* single); education (any qualification *v.* no educational qualifications (i.e., no O-levels, professional qualifications or higher)); financial difficulties (yes *v.* no); family size (0–4; 5 or more children); smoking *v.* non-smoking; critical partner relationship derived from the Family Adversity Index,⁹ (low affection and high aggression, physical or emotional cruelty, no partner social support *v.* not present); poor housing defects (a summary variable of three indicators: inadequacy; basic living; and defects/infestation present *v.* not present); crime (in trouble with police) or conviction of the mother or father (yes *v.* no); and psychopathology of the mother (affective disorder, suicide attempts *v.* none).

In addition, the child’s gender and whether or not they were born prematurely (before 37 weeks gestation) was also recorded.

Statistical analysis

ALSPAC cohort

Data were collected on standardised forms that were returned to the study centre and encoded for computer analysis using SPSS 12.0 on a PC. The data for each child were double entered, checked and cleaned before being combined with the main data-set for analysis. Current ALSPAC children’s prevalence of disruptive behaviour disorders were compared with never ALSPAC children as well as previous ALSPAC children diagnoses using categorical χ^2 tests (Question 1). Combining current and previous ALSPAC children provides an approximate estimate of the prevalence that would be found in the original ALSPAC cohort, excluding those who dropped out for whom we did not have teacher data. The prevalence of disruptive behaviour disorder in this ‘total ALSPAC’ group was then compared with that in the never ALSPAC group.

To determine whether participant drop-out was random or systematic, previous ALSPAC children were compared with current ALSPAC children on factors previously shown to predict disruptive behaviour problems (Question 2).^{10–12} Categorical outcomes were compared using χ^2 tests, and continuous outcomes with the use of Mann–Whitney tests for ordinal data. To determine the independent factors best predicting drop-out, all precursors were entered into multiple logistic regression (outcome: previous ALSPAC *v.* current ALSPAC) and individually adjusted for all other precursor variables. To answer whether prediction models are still valid despite participant drop-out, univariate logistic regressions were computed separately for the current ALSPAC and previous ALSPAC children employing factors previously reported to predict disruptive behaviour disorder (Question 3). The outcome was any disruptive behaviour disorder (ADHD and behaviour disorders combined) *v.* no disruptive behaviour diagnosis. Individual factors assessed in pregnancy and previously reported to predict disruptive behaviour disorder (i.e. male gender,¹³ prematurity,^{14,15} socioeconomic disadvantage,¹⁰ smoking in pregnancy,¹¹ critical partner relationship,^{16,17} parents’ previous crime involvement^{18,19} or maternal psychopathology²⁰ were entered as predictors of any disruptive behaviour disorder *v.* no disorder in separate univariate regression analyses for the current ALSPAC participants (260 with a positive diagnosis *v.* 3712 with no positive diagnosis) and previous ALSPAC participants (72 with a positive diagnosis *v.* 1058 with no positive diagnosis). To determine statistical difference in prediction, previous and current ALSPAC (factor: group membership) were combined and the interaction between group membership and individual predictor was computed. None of the interaction terms should be statistically significant if the prediction model did not differ between current and previous ALSPAC children.

Simulations

A series of 36 simulations was carried out to explore the impact of selective participant drop-out on the prediction of Y (disruptive behaviour) from a predictor X. Of primary interest were simulations in which drop out and disruptive behaviour were predicted by the same factor (X) (i.e. the drop-out occurred by selection on a predictor X in regression) and the degree of selection was varied between simulations. In each simulation, we generated a sample of 5000 cases, which was then subjected to a drop-out process. Each case *i* was characterised by a predictor value X_i and a criterion value Y_i , such that X and Y approximated a bivariate standard

normal distribution in the sample. The correlation between X and Y varied between simulations, in the range of 0.1–0.9, in steps of 0.1 (note that, because the variables were standardised, the Pearson correlation coefficient is identical to the linear regression coefficient in an ordinary least-squares model). For each correlation level, we simulated four stochastic drop-out processes, which differed in selectivity (although keeping the overall drop-out rate constant). We used the following drop-out rule:

$$\delta_i = \frac{1}{1 + \exp(-X_i/\tau)}$$

in which δ_i was the probability that case i was dropped from the sample, and τ was a scaling parameter that was manipulated between simulations. The general form of this logistic rule is shown in Fig. 2. For each value of τ , the expected proportion of dropped cases is 0.5. In all the simulations, the proportion of dropped cases was within the 0.49–0.51 range. The drop-out process was more selective (i.e. dependent on the value of X) for lower values of τ . Across a typical simulated 5000-case sample, the point-biserial correlations between X and a binary drop-out indicator were 0.10, 0.42, 0.61 and 0.78 for τ values of 5, 1, 0.5, and 0.1 respectively, confirming the high selectivity of drop-out for the lower values of τ .

A second set of simulations was carried out to determine the effect on the regression between X and Y (disruptive behaviour) of

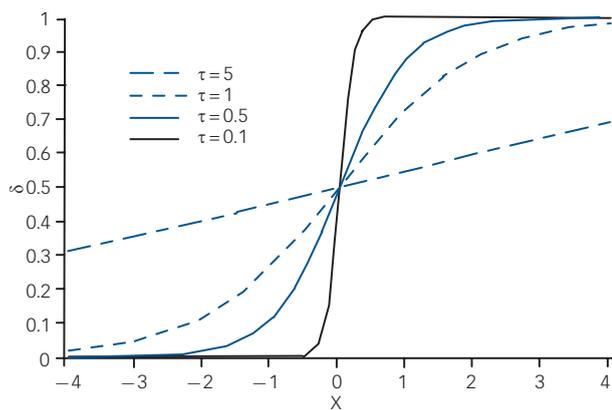


Fig. 2 Probability of dropping out (δ) as a function of X, for different values of τ .

drop-out that is selective on the criterion Y (e.g. drop-out of cases with higher scores on the criterion variable) and of drop-out that is selective on both the predictor (X) and the criterion variable (Y).

Results

Prevalence of disruptive behaviour disorder

As shown in Table 1 our total ALSPAC group had a lower prevalence of all teacher-based disorders than the unselected never ALSPAC group, although the findings in relation to any oppositional/conduct disorder ($P=0.075$) are marginal. This ‘prevalence gap’ might be explained by our missing data for some of those who dropped out and/or by selection bias that was operating even at initial recruitment. However, the prevalence of the total ALSPAC and never ALSPAC groups was closer than the current and never groups, suggesting that the initial cohort was more representative for teacher-reported disruptive behaviour disorder than after drop-out had occurred. Nevertheless, some selection had occurred over time according to the criterion, disruptive behaviour disorder.

Is drop-out selective or random?

The comparisons between the current and previous ALSPAC children are shown in Table 2. Drop-out from ALSPAC was systematically related to having a mother who was single, had no educational qualifications, encountered financial difficulties, being raised in a large family where the mother smoked, had a poor relationship with the partner, lived in poor housing, had been involved in crime and been convicted or suffered psychopathology during pregnancy. When prediction was adjusted for all other factors, being single (odds ratio (OR) = 1.45, 95% CI 1.19–1.77), family size (OR = 3.17, 95% CI 1.55–6.46), smoking (OR) = 1.41, 95% CI 1.15–1.73), no educational qualifications (OR) = 1.35, 95% CI 1.07–1.71) and financial difficulties (OR) = 1.39, 95% CI 1.07–1.81) remained significant independent predictors of drop-out.

Does drop-out reduce the validity of prediction of disruptive behaviour disorder?

Disruptive behaviour prediction with the ALSPAC data

The same variables that were related to the drop-out process were used as predictors for the disruptive behaviour disorder criterion. The individual predictors and the magnitude of prediction were very similar for the previous and current ALSPAC groups.

	Never ALSPAC	Current ALSPAC	Previous ALSPAC	Total ALSPAC	Test for 3-group differences ^a		Test for 2-group differences ^b	
	(<i>n</i> = 4383)	(<i>n</i> = 3946)	(<i>n</i> = 1130)	(<i>n</i> = 5076)	(never v. current v. previous)		(total v. never)	
	% (<i>n</i>)	χ^2	<i>P</i>	χ^2	<i>P</i>			
Any ADHD	3.8 (165)	2.4 (93)	4.8 (55)	2.9 (148)	22.6	<i>P</i> < 0.001	5.30	<i>P</i> = 0.021
Inattentive ADHD	1.6 (71)	1.2 (47)	2.3 (26)	1.4 (73)	7.74	<i>P</i> = 0.021	0.52	<i>P</i> = 0.472
Hyperactive ADHD	0.6 (26)	0.3 (12)	0.8 (9)	0.4 (21)	5.84	<i>P</i> = 0.054	1.53	<i>P</i> = 0.216
Combined ADHD	1.6 (68)	0.9 (34)	1.8 (20)	1.1 (54)	10.09	<i>P</i> = 0.006	4.39	<i>P</i> = 0.036
Any oppositional or conduct disorder	3.1 (138)	2.1 (84)	4.0 (45)	2.5 (129)	14.16	<i>P</i> = 0.001	3.16	<i>P</i> = 0.075
Oppositional defiant disorder	2.0 (86)	1.3 (52)	2.3 (26)	1.5 (78)	7.48	<i>P</i> = 0.024	2.50	<i>P</i> = 0.114
Conduct disorder	1.2 (52)	0.8 (32)	1.7 (19)	1.0 (51)	6.90	<i>P</i> = 0.032	0.72	<i>P</i> = 0.396
Any disruptive behaviour disorder	5.2 (228)	3.5 (139)	6.4 (72)	4.2 (211)	21.92	<i>P</i> < 0.001	5.81	<i>P</i> = 0.016

ADHD, attention-deficit hyperactivity disorder; ALSPAC, Avon Longitudinal Study of Parents and Children.
 a. Slight discrepancy in *n* due to availability of teacher diagnosis.
 b. Comparison: never ALSPAC, current ALSPAC and previous ALSPAC.
 c. Comparison: never ALSPAC and total ALSPAC.

Table 2 Prediction of drop-out (current v. previous ALSPAC participants)

	n	Prevalence, % ^a		Prediction of drop-out					
		Current	Previous	Unadjusted			Adjusted (n = 4070)		
				OR	95% CI	P	OR	95% CI	P
Child gender, male	5115	51.0	50.4	0.97	0.85–1.11	0.702	0.93	0.79–1.08	0.341
Born prematurely	5115	4.8	6.1	1.28	0.96–1.69	0.092	1.28	0.90–1.82	0.163
Marital status, single	4957	17.3	29.1	1.97	1.68–2.30	<0.001	1.45	1.19–1.77	<0.001
Education, no qualifications	4879	10.7	17.2	1.73	1.43–2.10	<0.001	1.35	1.07–1.71	0.011
Financial difficulties	4713	7.7	12.9	1.77	1.42–2.21	<0.001	1.39	1.07–1.81	0.015
Family size, >4 children	4984	0.7	2.6	3.58	2.12–6.04	<0.001	3.17	1.55–6.46	0.002
Maternal smoking	4452	15.3	25.1	1.85	1.55–2.21	<0.001	1.41	1.15–1.73	0.001
Critical partner relationship	5058	14.4	19.6	1.44	1.21–1.71	<0.001	1.08	0.86–1.34	0.512
Housing	5041	15.4	18.5	1.25	1.05–1.48	0.014	1.10	0.89–1.36	0.379
Crime and conviction	4547	1.7	3.5	2.12	1.37–3.26	0.001	1.14	0.68–1.90	0.620
Psychopathology of mother	4889	23.2	30.3	1.44	1.23–1.67	<0.001	1.18	0.98–1.42	0.081

a. Maximum n for current and previous ALSPAC is 3975 and 1140 respectively. Bonferoni adjusted P (0.05/11) = 0.0045.

Table 3 Simple univariable prediction of disruptive behaviour disorder for the current ALSPAC and previous ALSPAC children (those who have dropped out) using factors assessed during pregnancy

	n	Current ^a			Previous ^b			Interaction ^c			
		OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	
Child gender, male	5102	3.02	2.26	4.02	<0.001	4.40	2.43–7.99	<0.001	1.46	0.75–2.83	0.263
Born prematurely	5102	1.55	0.94	2.56	0.088	1.76	0.77–4.00	0.177	1.14	0.43–2.98	0.796
Marital status, single	4944	1.72	1.28	2.31	<0.001	1.89	1.14–3.15	0.014	1.10	0.61–1.98	0.746
Education, no qualifications	4867	1.53	1.07	2.19	0.021	2.30	1.30–4.07	0.004	1.51	0.77–2.96	0.233
Financial difficulties	4702	2.35	1.62	3.41	<0.001	2.71	1.47–4.97	0.001	1.15	0.56–2.35	0.697
Family size, >4 children	4971	1.08	0.26	4.56	0.917	1.82	0.54–6.20	0.336	1.69	0.25–11.21	0.586
Maternal smoking	4442	1.98	1.45	2.71	<0.001	2.52	1.44–4.41	0.001	1.27	0.67–2.42	0.459
Critical partner relationship	5045	1.89	1.39	2.55	<0.001	2.36	1.42–3.94	0.001	1.25	0.69–2.27	0.458
Housing	5028	1.32	0.96	1.83	0.090	1.40	0.78–2.50	0.261	1.06	0.54–2.06	0.874
Crime and conviction	4537	1.98	0.89	4.40	0.094	1.69	0.50–5.72	0.400	0.85	0.20–3.67	0.831
Psychopathology of mother	4877	2.31	1.77	3.02	<0.001	2.08	1.25–3.45	0.005	0.90	0.51–1.60	0.717

a. Two hundred and sixty individuals with disruptive behaviour disorder and 3712 controls.
 b. Seventy-two individuals with disruptive behaviour disorder and 1058 controls.
 c. Group (current v. previous ALSPAC). All children: 332 individuals with disruptive behaviour disorder and 4770 controls. Bonferoni adjusted P (0.05/11) = 0.0045.

Teacher-reported disruptive behaviour disorder in middle childhood was more likely when parents had low education, financial difficulties or critical partner relationships, when the mother had psychopathology or smoked in pregnancy, and for boys (Table 3). There were no significant interactions between group membership (previous ALSPAC v. current ALSPAC) and individual predictors (e.g. financial difficulties) when predicting the presence or absence of disruptive behaviour disorder, i.e. the same predictive model seemed to apply equally well to previous and current ALSPAC participants.

The simulations

Figure 3 gives an overview of the observed correlations between X and Y before and after the drop-out process in the simulations in which drop-out was selective on X. The results show that the drop-out process related to X has an effect on the correlations between X and Y. Figure 3 shows that in all simulations the correlation between X and Y was reduced in all simulations, and that the suppression effect was somewhat larger for the more selective drop-out processes (i.e. in those simulations in which τ was small).

Figure 4 demonstrates the effect of the drop-out process on a simulated sample. In this example, the correlation in the original

sample was high at $r = 0.90$, and the drop-out process was highly selective ($\tau = 0.1$). The plot shows that the variance in the sample was reduced on both predictor (X) and criterion variable (Y). However, the non-standardised slope of the best-fitting regression

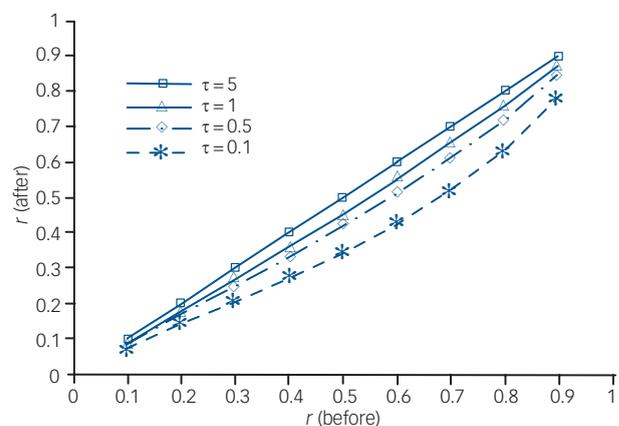


Fig. 3 Correlation between predictor X and criterion Y before and after drop-out, as a function of τ .

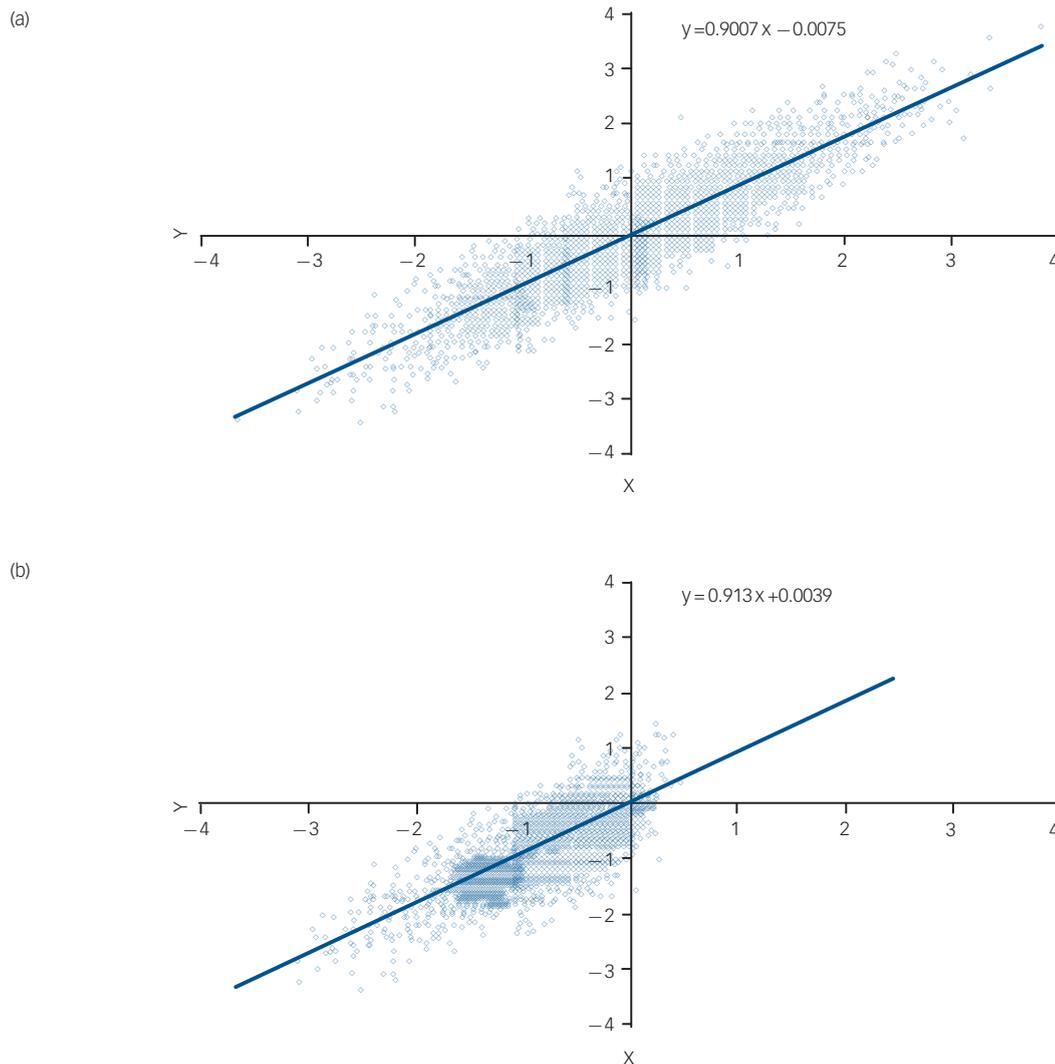


Fig. 4 Simulated effect of selective drop-out according to the predictor variable X on least-squares linear regression model. X = predictor, Y = criterion. (a) before drop-out and (b) after drop-out.

line was practically unaltered by the drop-out process. The correlation (which corresponds to the standardised regression coefficient), was reduced from 0.90 to 0.78 after drop-out, as can be seen in Fig. 4.

The simulations demonstrate that selection on X in a regression has the effect of reducing the variance in X (and Y) and attenuates the correlation between X and Y. As shown here, the effects of selective drop-out in X on predictor–criterion correlation (and, by implication, regression) can be relatively small, even under a highly selective drop-out regime. Range restriction as a result of selective drop-out does not necessarily affect the validity of a regression model, although it can lead to underestimation of the criterion–predictor correlation.

It is important to note that drop-out by selection on the criterion variable (Y) can have a very different effect on the regression coefficients. Figure 5(a) shows an example, based on the same original simulated sample as in Fig. 4 ($r = 0.90$ before drop-out, $\tau = 0.1$). Figure 5(a) shows that both the regression and correlation coefficient were reduced as a result of drop-out of participants with higher scores on the criterion variable ($r = 0.79$ after drop-out). Figure 5(b) provides an example in which there was selective drop-out on both the predictor (X) and the criterion variable (Y),

with participants that scored highly on both variables more likely to drop out. Drop-out that was selective on both variables suppressed the regression coefficient (but less so than in the example in which drop-out was selective on the criterion only) and also reduced the correlation between predictor and criterion ($r = 0.77$ after drop-out).

Discussion

We examined whether those who continued to participate in a longitudinal study of disruptive behaviour disorders differed from those who previously were enrolled but dropped out. To allow for comparisons of prevalence and to test whether longitudinal prediction is affected by drop-out, as often claimed in textbooks,^{21,22} the outcome was the presence of a diagnosis of a disruptive behavioural disorder based on teacher reports.

Selective drop-out and prevalence

Drop-out was considerable, with teacher returns on 37% (5115/13 971) of those believed to be alive or 49% (5119/10 431) of those eligible to be contacted. We only consider here the response to one

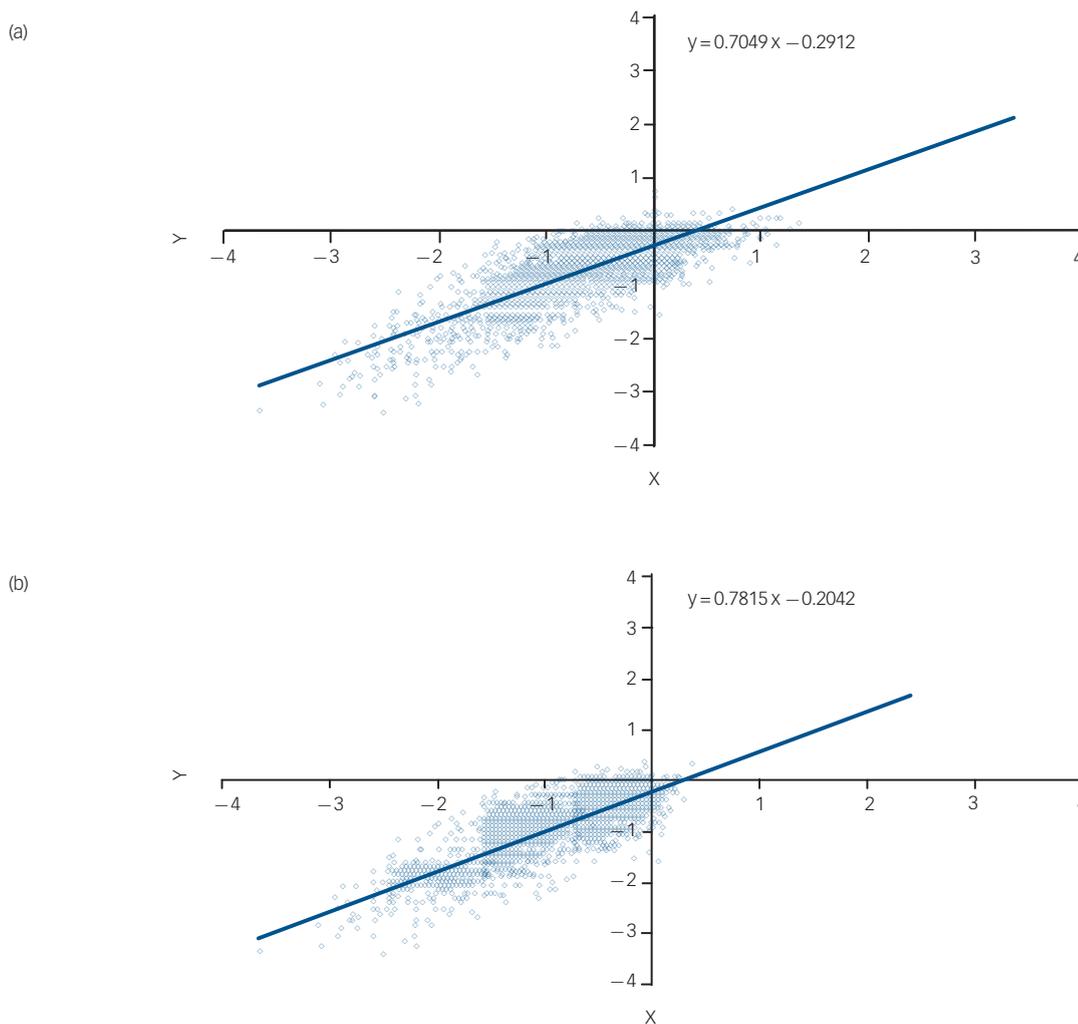


Fig. 5 Simulated effect of selective drop-out (a) after drop-out according to the criterion variable *y* on least-squares linear regression model and (b) after drop-out according to the predictor variable *x* and criterion variable *y* on least-squares linear regression model. *X* = predictor; *Y* = criterion.

particular assessment during the eighth year of life of the child. The participation rate is higher for any contact in a given year, whether for face-to-face assessments or other questionnaires.⁵ Overall, the follow-up rate is similar to recent comparable large-scale longitudinal studies with repeated assessments.^{23,24} In general, participation rates are higher in older cohorts enrolled some decades ago,^{25,26} for studies focused on specific high-risk samples in the first place^{2,27} or for samples that were small and selective.²⁸

The attrition from the sample we studied was systematically related to family characteristics, which supports the conclusions of previous work^{3,4,27} that psychosocial factors are associated with attrition in longitudinal studies. The selective drop-out of participants had an impact on the prevalence of teacher-reported disruptive behaviour disorders, with the prevalence among children who were still participating being approximately half that of children who had dropped out. The factors that influenced retention in the ALSPAC sample also influenced the likelihood of disruptive behaviour disorder, i.e. the missingness was non-ignorable.²⁹ Longitudinal studies are likely to underestimate the prevalence and incidence of disorders as shown here and elsewhere.³⁰ Cross-sectional studies requiring only one single assessment are likely to be a more accurate in estimating prevalence.³¹

Selective drop-out and prediction

Finally, we investigated whether selective drop-out of participants does reduce the validity of prediction from longitudinal analysis. Prospective studies can only rely on the data of the individuals who continue to participate or they have to estimate missing data using sophisticated missing value substitution modelling and imputations.^{32,33} To our knowledge, this is the first investigation that could compare the prediction of outcomes of current and previous participants in a prospective study. We found that selective drop-out of participants according to a range of predictor variables did not invalidate the prediction of teacher-reported disruptive behaviour disorders by factors that were assessed as early as pregnancy and birth that have previously been shown to predict these difficulties.¹⁰⁻¹² Boys were significantly more likely to develop teacher-reported disruptive behaviour disorder, as were the children of mothers who suffered psychopathology or smoked during pregnancy, who had poor partner relationships or who were single, poorly educated or suffered financial hardship.³⁴ These same predictions were found for those who were still participating in the ALSPAC study as well as for those who had dropped out. Despite reduction to a super-normal current ALSPAC sample, the predictive factors and their strength were

about the same as for the previous participants. Contrary to common assumptions,^{21,35} the presence of a substantial selection bias did not markedly attenuate the relationship between exposure and outcome in this study. Although prevalence rates do have an impact on statistical power, differences in prevalence *per se* did not alter prediction in this instance. Similarly, Moffitt and colleagues¹³ investigated factors suspected to predict disruptive behaviour disorders in a sample of approximately 1000 children, half of them girls. They found, that despite girls being much less likely to develop disruptive behaviour disorder (low prevalence), the same factors predicted disruptive behaviour problems in both girls and boys.

We conducted simulations to explain why the effects of selective participant drop-out on predictor–criterion correlation (and regression) were relatively small in our empirical study. We found that a range of social and parental variables previously described as precursors of disruptive behaviour disorder in children affected the drop-out process. The simulations confirmed that if the selection is on X in a regression, the effect is one of reducing the variance in X (and Y), not affecting the regression but attenuating the correlation between X and Y (see Berk,³⁶ p. 389). Our simulations add that even under a highly selective drop-out regime related to X, the overall reduction in the correlations is small to moderate (Fig. 3). Therefore, range restriction as a result of selective drop-out according to X does not affect the internal or external validity of the regression model,³⁶ although the correlation coefficient after selective drop-out may underestimate the true correlation between the predictor and criterion variable.

In our empirical ALSPAC study, we see little evidence that teachers selectively underreported on children with disruptive behaviour. It seems unlikely that teachers would have been less likely to report on those with more disruptive behaviour since teachers are usually well aware of those who disturb lessons.³⁷ Nevertheless, we carried out a second set of simulations (examples shown in Fig. 5(a) and (b)) that showed that if selection on the criterion (Y) had occurred (i.e. if those with high disruptive behaviour disorder were less likely included in the sample), then the regression would be attenuated and the original regression line would no longer fit the data. Although confirming that the internal and external validity are weakened in these circumstances³⁶ and the true relationship between X and Y is systematically underestimated, our simulation also demonstrated that when drop-out is influenced by the predictor as well as the criterion variable, this only mildly reduces estimates of the slope of the true regression line.

We conclude that the regression coefficients hold for the current, previous and entire cohort due to the fact that, despite selection bias on X (and thus restricted range), the differences between the current and previous groups with disruptive behaviour disorder are small. Where the predictor variables have small to moderate (linear) associations with both, the drop-out of participants and the outcome variable, the impact on the predictor–criterion regression is small. However, if the drop-out process is dependent on the criterion variable (e.g. high scorers systematically excluded), then internal and external validity is threatened and the true relationship between predictor and criterion can no longer be estimated reliably. Particularly in cases where the selection process follows a complex pattern (e.g. with dependencies on several variables or non-monotonic dependencies; see Berk³⁶ for a full discussion) internal and external validity are under threat.

Limitations and conclusions

There are limitations to our study. Even fewer teachers than parents completed the diagnostic instrument in the current

sample and this itself could have introduced bias. For example, teachers may have been more likely to complete the DAWBA in well-organised affluent schools. However, as these schools are also likely to have a lower prevalence of disruptive behaviour disorders, and possibly better strategies of managing them, this would be likely to lower prevalence across all three groups. Teachers would not have been aware of which children were or had been participants in ALSPAC when they completed measures on all the children in their class. Our diagnosis of a disruptive behaviour disorder was based only on teacher reports, although the limitation of having just one informant is partly offset by the fact that teacher reports are particularly informative for diagnoses for externalising disorders.³⁷ Nevertheless, our findings may not be applicable to diagnoses of a disruptive behaviour disorder based on parent data, self-report data or multi-informant data, or indeed to other outcomes within this or other studies.

In conclusion, participant loss in the ALSPAC cohort was systematic, with children with teacher-reported disruptive behaviour disorder being more frequently lost to follow-up. Our results suggest that longitudinal studies are likely to underestimate the prevalence and incidence of disorders,⁴ but that this might not negate findings in relation to the predictors of disorder if selection occurs according to the predictor variables. Our results need replication in relation to other cohorts and other outcomes. However, the simulations indicate that despite highly selective drop-out as a result of X and reduced range in both predictor and outcome variables, the regression parameter estimates are only mildly affected. Our demonstrations do not imply that selective drop-out is always harmless. For instance, selective drop-out effects can have significant implications if the selection is according to the outcome variable, if the drop-out process is complex or incidental³⁶ or there is a non-linear relation between predictor(s) and criterion. In such cases, explicit modelling of the drop-out process (e.g. Diggle & Kenward³⁸ and Little³⁹) might help to clarify the implications of drop-out for model validity. Nevertheless, although everything should be done to reduce participant loss in cohort studies,^{40,41} it is reassuring to find that aetiological models from longitudinal samples can be valid and robust under specific conditions of selective loss of participants.

Dieter Wolke, PhD, Department of Psychology and Health Science Research Institute, Warwick Medical School, University of Warwick, Coventry; **Andrea Waylen**, PhD, Department of Oral and Dental Science, University of Bristol; **Muthanna Samara**, PhD, Department of Psychology, University of Warwick, Coventry; **Colin Steer**, MSc, Department of Community-based Medicine, University of Bristol; **Robert Goodman**, MD, Department of Child and Adolescent Psychiatry, Institute of Psychiatry, King's College London; **Tamsin Ford**, PhD, Institute of Health Service Research, Peninsula College of Medicine and Dentistry, Exeter; **Koen Lamberts**, PhD, Department of Psychology, University of Warwick, Coventry, UK

Correspondence: Dieter Wolke, Department of Psychology, University of Warwick, Coventry CV4 7AL, UK. Email: D.Wolke@warwick.ac.uk

First received 14 Apr 2008, final revision 7 Nov 2008, accepted 8 Jan 2009

Funding

The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC. This research was specifically funded by the Health Foundation to D.W., R.G., Jean Golding and Mike Beveridge (Grant 265/1981).

Acknowledgements

We are grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

References

- 1 Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med* 2008; **5**: e67.
- 2 Wolke D, Meyer R. Cognitive status, language attainment and pre-reading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. *Dev Med Child Neurol* 1999; **41**: 94–109.
- 3 Aylward G, Pfeiffer S. Follow-up and outcome of low birthweight infants: conceptual issues and a methodology review. *Aust Paediatr J* 1989; **25**: 3–5.
- 4 Wolke D, Söhne B, Ohrt B, Riegel K. Follow-up of preterm children: important to document dropouts. *Lancet* 1995; **345**: 447.
- 5 Golding J, Pembrey M, Jones R. ALSPAC – the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* 2001; **15**: 74–87.
- 6 Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psych Psychiatry* 2000; **41**: 645–55.
- 7 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th edn) text revision (DSM-IV-TR)*. APA, 2000.
- 8 Achenbach TM, Rescorla LA. Developmental issues in assessment, taxonomy, and diagnosis of psychopathology: life span and multicultural perspectives. In *Developmental Psychopathology Volume 1: Theory and Method* (eds D Cichetti, DJ Cohen): 139–80. John Wiley & Sons, 2006.
- 9 Bowen E, Heron J, Waylen A, Wolke D. Domestic violence risk during and after pregnancy: findings from a British longitudinal study. *BJOG* 2005; **112**: 1083–9.
- 10 Counts CA, Nigg JT, Stawicki JA, Rappley MD, von Eye A. Family adversity in DSM-IV ADHD combined and inattentive subtypes and associated disruptive behavior problems. *J Am Acad Child Adolesc Psychiatry* 2005; **44**: 690–8.
- 11 Kotimaa AJ, Moilanen I, Taanila A, Ebeling H, Smalley SL, McGough JJ, et al. Maternal smoking and hyperactivity in 8-year-old children. *J Am Acad Child Adolesc Psychiatry* 2003; **42**: 826–33.
- 12 Linnet KM, Dalsgaard S, Obel C, Wisborg K, Henriksen TB, Rodriguez A, et al. Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: review of the current evidence. *Am J Psychiatry* 2003; **160**: 1028–40.
- 13 Moffitt TE, Caspi A, Rutter M, Silva PA. *Sex Differences in Antisocial Behaviour*. Cambridge University Press, 2001.
- 14 Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002; **288**: 728–37.
- 15 Wolke D. The psychological development of prematurely born children. *Arch Dis Child* 1998; **78**: 567–70.
- 16 Johnston C, Mash EJ. Families of children with attention-deficit/hyperactivity disorder: review and recommendations for future research. *Clin Child Fam Psychol Rev* 2001; **4**: 183–207.
- 17 Fergusson D, Horwood L, Ridder E. Partner violence and mental health outcomes in a New Zealand birth cohort. *J Marriage Fam* 2005; **67**: 1103–19.
- 18 Farrington DP. The development of offending and antisocial behaviour from childhood: key findings from the Cambridge study in delinquent development. *J Child Psychol Psychiatry* 1995; **360**: 929–64.
- 19 Henry B, Caspi A, Moffitt TE, Silva PA. Temperamental and familial predictors of violent and non-violent criminal convictions: age 3 to age 18. *Dev Psychol* 1996; **32**: 614–23.
- 20 Cunningham CE, Boyle MH. Preschoolers at risk for attention-deficit hyperactivity disorder and oppositional defiant disorder: family, parenting, and behavioral correlates. *J Abnorm Child Psychol* 2002; **30**: 555–69.
- 21 Rothman K, Greenland S. *Modern Epidemiology*. Lippincott-Raven, 1998.
- 22 Szklo M, Nieto F. *Epidemiology*. Aspen, 2000.
- 23 Curtin T, Ingels S, Wu S, Heuer R, Owings J. *National Education Longitudinal Study of 1988: Base-Year to Fourth Follow-up Data File User's Manual*. US Department of Education, National Center for Education Statistics, 2002.
- 24 National Longitudinal Survey of Children and Youth. *National Longitudinal Survey of Children and Youth (NLSCY) – Overview Report*. Publications Centre Human Resources Development Canada, 1996.
- 25 Ferri E. Forty years on: Professor Neville Butler and the British Birth Cohort studies. *Paediatr Perinat Epidemiol* 1998; **12**: 31–44.
- 26 Fergusson DM, Boden JM, Horwood LJ. Examining the intergenerational transmission of violence in a New Zealand birth cohort. *Child Abuse Negl* 2006; **30**: 89–108.
- 27 Marlow N, Wolke D, Bracewell MA, Samara M. Neurologic and developmental disability at 6 years of age after extremely preterm birth. *N Engl J Med* 2005; **352**: 9–19.
- 28 Laucht M, Esser G, Baving L, Gerhold M, Hoesch I, Ihle W, et al. Behavioral sequelae of perinatal insults and early family adversity at 8 years of age. *J Am Acad Child Adolesc Psychiatry* 2000; **39**: 1229–37.
- 29 Parzen M, Lipsitz SR, Fitzmaurice GM, Ibrahim JG, Troxel A. Pseudo-likelihood methods for longitudinal binary data with non-ignorable missing responses and covariates. *Stat Med* 2005; **25**: 2784–96.
- 30 Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Arch Gen Psychiatry* 2003; **60**: 837–44.
- 31 Meltzer H, Gatward R, Goodman R, Ford T. *The Mental Health of Children and Adolescents in Great Britain: Summary Report*. TSO (The Stationery Office), 2000.
- 32 O'Hara Hines RJ, Hines WG. An appraisal of methods for the analysis of longitudinal categorical data with MAR drop-outs. *Stat Med* 2005; **24**: 3549–63.
- 33 Royston P. Multiple imputation of missing values: update of ice. *Stata J* 2005; **5**: 527–36.
- 34 Moffitt TE, Avshalom C. Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Dev Psychopathol* 2001; **13**: 355–75.
- 35 Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15**: 615–25.
- 36 Berk RA. An introduction to sample selection bias in sociological data. *Am Sociol Rev* 1983; **48**: 386–98.
- 37 Henry DB. Associations between peer nominations, teacher ratings, self-reports, and observations of malicious and disruptive behavior. *Assessment* 2006; **13**: 241–52.
- 38 Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. *Appl Stat* 1994; **43**: 49–93.
- 39 Little RJA. Modelling the drop-out mechanism in repeated-measures studies. *J Am Stat Assoc* 1995; **431**: 1112–21.
- 40 Farrington D, Gallagher B, Morley L, St. Ledger R, West D. Minimizing attrition in longitudinal research: methods of tracing and securing cooperation in a 24-year follow-up study. In *Data Quality in Longitudinal Research* (eds D Magnusson, L Bergman): 122–47. Cambridge University Press, 1990.
- 41 Ribisl KM, Walton MA, Mowbray CT, Luke DA, Davidson WS, Bootsmiller BJ. Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: Review and recommendations. *Eval Program Plann* 1996; **19**: 1–25.

