

University of Warwick institutional repository

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Authors:	Dieter Wolke, Andrea Waylen, Muthanna Samara, Colin Steer, Robert Goodman, Tamsin Ford and Koen Lamberts
Title:	Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders
Year of publication:	2009
Link to published version:	<a href="http://dx.doi.org/10.1192/bjp.bp.108.053751">http://dx.doi.org/10.1192/bjp.bp.108.053751</a>
Publisher statement:	This is an author-produced electronic version of an article accepted for publication in the British Journal of Psychiatry. The definitive publisher-authenticated version is available online at <a href="http://bjp.rcpsych.org">http://bjp.rcpsych.org</a>



**Does selective dropout in longitudinal studies lead to biased prediction of behaviour disorders?**

Dieter Wolke, Ph.D.\*; Andrea Waylen, Ph.D. †; Muthanna Samara, Ph.D.\*;

Colin Steer, M.Sc. ‡; Robert Goodman, M.D.§ ,Tamsin Ford, Ph.D.+ and Koen Lamberts, Ph.D.\*

From the \* University of Warwick, Department of Psychology and Health Science Research Institute, Warwick Medical School, Coventry (UK); † University of Bristol, Department of Oral and Dental Science, Bristol (UK); ‡ University of Bristol, Department of Community-based Medicine, Bristol (UK); § King's College London Institute of Psychiatry, Department of Child and Adolescent Psychiatry, London (UK). + Peninsula College of Medicine and Dentistry, Exeter, (UK)

Corresponding Author: Prof. Dieter Wolke, University of Warwick, Department of Psychology, Coventry CV4 7AL, Tel. 0044 24 7652 3537; Fax: 0044 24 7652 4225; email: D.Wolke@warwick.ac.uk

**ACKNOWLEDGEMENT**

We are extremely 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. 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 Dieter Wolke, Jean Golding, Robert Goodman & Mike Beveridge (Grant 265/1981)

**Running Head:** Selective dropout in longitudinal studies

## **ABSTRACT**

### **Background**

Participant dropout 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 dropout 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 dropout 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 dropout 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.

### **Conclusion**

Cohort studies are prone to selective dropout 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 dropout.

**Key words:** subject loss, prediction, ADHD, Conduct disorder, oppositional defiant disorder, ALSPAC

**Declarations of interest:** none

## INTRODUCTION

Prospective studies provide one of the strongest methodologies for studying aetiological mechanisms<sup>1</sup>, but are vulnerable to selection biases due to losses to follow up. Subject loss can be random<sup>2</sup> or systematically related to social or biological characteristics of the participants that may or may not be associated with the outcome of interest<sup>3 4</sup>. 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 dropout 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. Firstly, 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**)? Secondly, do those who have dropped out of the cohort (**previous ALSPAC**) differ systematically from those who stayed on (**current ALSPAC**)? Thirdly, are the prediction models for disruptive behaviour disorders the same for those who are currently still participating in the study (**current ALSPAC**) compared to those who dropped out (**previous ALSPAC**)? Finally, we conducted simulations to explain the impact of selective dropout on the strength of prediction if dropout, predictor and criterion variables are correlated to varying degrees.

## METHODS

### Participants

ALSPAC<sup>5</sup> 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. 13,988 infants were alive at one year and 13,971 at 7 years of age. When compared to 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 minorities<sup>5</sup>.

(Figure 1 here)

At 7 years 9 months, as part of a study on disruptive behaviour disorders (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)<sup>6</sup> on all the children in their class with a birth date between April 1991 and December 1992. From a total of 10431 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) (see Fig. 1). The teacher completion was thus 5115/10431 of eligible children (49%) or 5115/13971 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 self-report questionnaire with regard to medication, symptoms, diet and lifestyle, attitudes and behaviour, and social-environmental features<sup>5</sup>. 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<sup>6</sup>) 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<sup>7</sup>, 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<sup>8</sup>.

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), while 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 vs. single); education (any qualification vs. no educational qualifications (i.e., no O-levels, professional qualifications or higher); financial difficulties (yes vs no); family size (0-4; 5 or more children), smoking vs non-smoking, critical partner relationship derived from the FAI (<sup>9</sup> (low affection and high aggression, physical or emotional cruelty, no partner social support vs not present), poor housing defects (a summary variable of 3 indicators: a) inadequacy, b) basic living, c) defects/infestation present vs. not present), crime (in trouble with police) or conviction of the mother or father (yes vs. no) and psychopathology of the mother (affective disorder, suicide attempts vs none). In addition, the child's sex and whether or not s/he was born prematurely (before 37 weeks gestation) were 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. 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 to never ALSPAC children as well as previous ALSPAC children diagnoses using categorical  $\chi^2$  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 "drop outs" for whom we did not have teacher data. The prevalence of disruptive behaviour disorder in this "total ALSPAC" group was then compared to that in the never ALSPAC group.

To determine whether dropout was random or systematic, previous ALSPAC children were compared to current ALSPAC children on factors previously shown to predict disruptive behaviour problems (Question 2) <sup>10</sup> <sup>11</sup> <sup>12</sup>. Categorical outcomes were compared using  $\chi^2$  tests, and continuous outcomes with the use of Mann-Whitney tests for ordinal data. To determine the independent factors best predicting dropout, all precursors were entered into multiple logistic regression (outcome: previous ALSPAC vs current ALSPAC) and individually adjusted for all other precursor variables. To answer whether prediction models are still valid despite subject dropout, 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) vs. no disruptive behaviour diagnosis. Individual factors assessed in pregnancy and previously reported to

predict disruptive behaviour disorder (i.e., male sex<sup>13</sup>, prematurity<sup>14 15</sup>, socio-economic disadvantage<sup>10</sup> smoking in pregnancy<sup>11</sup>, critical partner relationship<sup>16 17</sup> parents' previous crime involvement<sup>18 19</sup> or maternal psychopathology<sup>20</sup> were entered as predictors of any disruptive behaviour disorder vs. no disorder in separate univariate regression analyses for the current ALSPAC participants (260 cases vs. 3712 with no positive diagnosis) and previous ALSPAC participants (72 cases vs 1058). 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 dropout on the prediction of Y (disruptive behaviour) from a predictor X. Of primary interest were simulations in which dropout and disruptive behaviour were predicted by the same factor (X), to varying degree (i.e., in which dropout occurred by selection on a predictor in regression) (Dieter, the previous sentence is not clear). In each simulation, we generated a sample of 5000 cases, which was then subjected to a dropout process. Each case  $i$  was characterized 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 from 0.1 to 0.9, in steps of 0.1 (note that, because the variables were standardized, 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 dropout processes, which differed in selectivity (while keeping the overall dropout rate constant). We used the following dropout rule:

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

in which  $\delta_i$  was the probability that case  $i$  was dropped from the sample, and  $\tau$  was a scaling parameter that was manipulated between simulations. The general form of this logistic rule is shown in Figure 2. For each value of  $\tau$ , the expected proportion of dropped cases is 0.5. In all the simulations, the proportion of dropped cases was within the .49 to .51 range. The dropout process was more selective (i.e., dependent on the value of X) for lower values of  $\tau$ . Across a typical simulated 5000-case sample, the point-biserial correlations between X and a binary dropout indicator were .10, .42, .61 and .78, for  $\tau$  values of 5, 1, 0.5, and 0.1, respectively, confirming the high selectivity of dropout for the lower values of  $\tau$ .



(Figure 2 here)

A second set of simulations was carried out to determine the effect on the regression between X and Y (disruptive behaviour) of dropout that is selective on the criterion Y (e.g., dropout of cases with higher scores on the criterion variable) and of dropout that is selective on both the predictor (X) and the criterion variable (Y).

## RESULTS

### **Does the prevalence of disruptive behaviour disorder differ between current, previous and never ALSPAC children?**

(Table 1 here)

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 ( $p=.075$ ) are marginal. This “prevalence gap” might be explained by our missing data on some dropouts 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 dropout had occurred. Nevertheless, some selection had occurred over time according to the criterion, disruptive behaviour disorder.

### **Is Dropout Selective or Random?**

The comparisons between the current and previous ALSPAC children are shown in table 2. Dropout 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 (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 dropout.

(Table 2 here)

### **Does dropout reduce the validity of prediction of disruptive behaviour disorder?**

#### Disruptive Behaviour Prediction with the ALSPAC Data

The same variables that were related to the dropout 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. 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 (see table 3). There were no significant interactions between group membership (previous ALSPAC vs. 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.

(Table 3 here)

#### The Simulations

Figure 3 gives an overview of the observed correlations between X and Y before and after the dropout process in the *simulations* in which dropout was selective on X. The results show that the dropout 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 dropout processes (i.e., in those simulations in which  $\tau$  was small).

(Figure 3 here)

Figure 4 demonstrates the effect of the dropout process on a simulated sample. In this example, the correlation in the original sample was high at  $r = .90$ , and the dropout 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-standardized slope of the best-fitting regression line was practically unaltered by the dropout process. The correlation (which corresponds to the standardized regression coefficient), was reduced from .90 to .78 after dropout, as can be seen in Figure 4.

(Figure 4 here)

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 dropout in X on predictor-criterion correlation (and, by implication, regression) can be relatively small, even under a highly selective dropout regime. Range restriction as a result of selective dropout 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 dropout by selection on the criterion variable (Y) can have a very different effect on the regression coefficients. Figure 5a shows an example, based on the same original simulated sample as in Figure 4 ( $r = .90$  before dropout;  $\tau = 0.1$ ). Figure 5a shows that both the regression and correlation coefficient were reduced as a result of dropout of cases with higher scores on the criterion variable ( $r = .79$  after dropout). Figure 5b provides an example in which there was selective dropout on both the predictor (X) and the criterion variable (Y), with cases that scored highly on both variables more likely to drop out. Dropout that was selective on both variables suppressed the regression coefficient (but less so than in the example in which dropout was selective on the criterion only), and also reduced the correlation between predictor and criterion ( $r = .77$  after dropout).

## 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 dropout, as often claimed in textbooks<sup>21</sup><sup>22</sup>, the outcome was the presence of a diagnosis of a disruptive behavioural disorder based on teacher reports.

Dropout was considerable, with teacher returns on 37% (5115/13971) of those believed to be alive or 49% (5119/10431) of those eligible to be contacted. We only consider here the response to one particular assessment during the 8<sup>th</sup> 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<sup>5</sup>. Overall, the follow-up rate is similar to recent comparable large scale longitudinal studies with repeated assessments<sup>23 24</sup>. In general, participation rates are higher in older cohorts enrolled some decades ago<sup>25 26</sup>, for studies focussed on specific high risk samples in the first place<sup>27 28</sup> or for samples that were small and selective.<sup>29</sup>

The attrition from the sample we studied was systematically related to family characteristics, which supports the conclusions of previous work<sup>3 4 28</sup> that psychosocial factors are associated with attrition in longitudinal studies. The selective dropout 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<sup>30</sup>. Longitudinal studies are likely to underestimate the prevalence and incidence of disorders as shown here and elsewhere<sup>31</sup>. Cross-sectional studies requiring only one single assessment are likely to be a more accurate in estimating prevalence<sup>32</sup>.

Finally, we investigated whether selective dropout does reduce the validity of prediction from longitudinal analysis. Prospective studies can only rely on the data of the subjects who continue to participate or they have to estimate missing data using sophisticated missing value substitution modelling and imputations<sup>33 34</sup>. 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 dropout 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<sup>10 11 12</sup>. 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<sup>35</sup>. 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,<sup>36 21</sup> the presence of

a substantial selection bias did not markedly attenuate the relationship between exposure and outcome in this study. While prevalence rates do have an impact on statistical power, differences in prevalence per se did not alter prediction in this instance. Similarly, Moffitt and colleagues<sup>13</sup> 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 dropout 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 dropout 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<sup>37</sup> (p. 389). Our simulations add that even under a highly selective dropout regime related to X, the overall reduction in the correlations is small to moderate (Figure 3). Therefore, range restriction due to selective dropout according to X does not affect the internal or external validity of the regression model<sup>37</sup> although the correlation coefficient after selective dropout may underestimate the true correlation between the predictor and criterion variable.

In our empirical ALSPAC study, we see little evidence that teachers selectively under-reported 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<sup>38</sup>. Nevertheless, we carried out a second set of simulations (examples shown in Figure 5a 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. While confirming that the internal and external validity are weakened in these circumstances<sup>37</sup> and the true relationship between X and Y is systematically underestimated, our simulation also demonstrated that when dropout 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 dropout and the outcome variable, the impact on the predictor-criterion

regression is small. However, if the dropout 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 <sup>37</sup> for a full discussion) internal and external validity are under threat.

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, though the limitation of having just one informant is partly offset by the fact that teacher-reports are particularly informative for diagnoses for externalizing disorders <sup>38</sup>. 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 <sup>4</sup>, 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 dropout due to 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 dropout is always harmless. For instance, selective dropout effects can have significant implications if the selection is according to the outcome variable, if the dropout process is complex or incidental <sup>37</sup> or there is a nonlinear relation between predictor(s) and criterion. In such cases, explicit modelling of the dropout process (e.g. <sup>39,40</sup>) might help to clarify the implications of dropout for model validity. Nevertheless, while everything should be done to reduce participant loss in cohort studies <sup>41 42</sup>, it is reassuring to find that aetiological models from longitudinal samples can be valid and robust under specific conditions of selective subject loss.

## REFERENCES

1. Vandenberg JP. Observational Research, Randomised Trials, and Two Views of Medical Science. *PLoS Med.* 2008;5(3):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(2):94-109.
3. Aylward G, Pfeiffer S. Follow-up and outcome of low birthweight infants: Conceptual issues and a methodology review. *Aust Paediat 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 (No. 8947):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(1):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(5):645-655.
7. Association AP. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association; 2000.
8. Achenbach TM, Rescorla LA. Developmental issues in assessment , taxonomy, and diagnosis of psychopathology: life span and multicultural perspectives. In: Cichetti D, Cohen DJ, eds. *Developmental Psychopathology Volume 1: Theory and Method*. Hoboken, N.J.: John Wiley & Sons; 2006:139-180.
9. Bowen E, Heron J, Waylen A, Wolke D. Domestic violence risk during and after pregnancy: findings from a British longitudinal study. *BJOG.* 2005;112:1083-1089.
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(7):690-698.
11. Kotimaa AJ, Moilanen I, Taanila A, Ebeling H, Smalley SL, McGough JJ, Hartikainen AL, Jarvelin MR. Maternal smoking and hyperactivity in 8-year-old children. *J Am Acad Child Adolesc Psychiatry* 2003;42(7):826-833.
12. Linnet KM, Dalsgaard S, Obel C, Wisborg K, Henriksen TB, Rodriguez A, Kotimaa A, Moilanen I, Thomsen PH, Olsen J, Jarvelin MR. Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: review of the current evidence. *Am J Psychiatry* 2003;160(6):1028-1040.
13. Moffitt TE, Caspi A, Rutter M, Silva PA. *Sex differences in antisocial behaviour*. Cambridge: 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(6):728-737.
15. Wolke D. The psychological development of prematurely born children. *Arch Dis Child.* 1998;78:567-570.
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(3):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(5):1103-1119.
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;36(6):929-964.
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;32:614-623.
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(6):555-569.
21. Rothman K, Greenland S. *Modern Epidemiology*. Philadelphia: Lippincott-Raven; 1998.
22. Szklo M, Nieto F. *Epidemiology*. Gaithersburg, MD: 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*. Washington, DC: U.S. Department of Education - National Center for Education Statistics; 2002.
24. NLSCY. *National Longitudinal Survey of Children and Youth (NLSCY) - Overview report*. Quebec: 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(2):89-108.
27. 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.
28. 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(1):9-19.
29. Laucht M, Esser G, Baving L, Gerhold M, Hoesch I, Ihle W, Steigleider P, Stock B, Stoehr RM, Weindrich D, Schmidt MH. Behavioral sequelae of perinatal insults and early family adversity at 8 years of age. *J Am Acad Child Adolesc Psychiatry* 2000;39(10):1229-1237.
30. 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;25:2784-2796.
31. 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(8):837-844.
32. Meltzer H, Gatward R, Goodman R, Ford T. *Mental health of children and adolescents in Great Britain.* London: The Stationery Office; 2000.
33. 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(23):3549-3563.
34. Royston P. Multiple imputation of missing values: Update of ice. *Stata Journal* 2005;5(4):527-536.
35. Moffitt TE, Avshalom C. Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Dev Psychopathol* 2001;13:355-375.
36. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15(5):615-625.
37. Berk RA. An Introduction to Sample Selection Bias in Sociological Data. *Am Sociol Rev.* 1983;48(3):386-398.
38. Henry DB. Associations between Peer Nominations, Teacher Ratings, Self-Reports, and Observations of Malicious and Disruptive Behavior. *Assessment* 2006;13(3):241-252.
39. Diggle P, Kenward MG. Informative Drop-Out in Longitudinal Data Analysis. *Appl Stat.* 1994;43:49-93.
40. Little RJA. Modelling the drop-out mechanism in repeated-measures studies. *J Am Stat Assoc.* 1995;431:1112-1121.
41. 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: Magnusson D, Bergman L, eds. *Data quality in longitudinal research.* Cambridge: Cambridge University Press; 1990:122-147.
42. 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):1-25.



**Table 1: Prevalence of disruptive behaviour disorder diagnoses according to cohort** (Slight N discrepancy due to availability of teacher diagnoses)

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

\* Comparison: never ALSPAC, current ALSPAC and previous ALSPAC

† Comparison: never ALSPAC and total ALSPAC

‡ Children co-morbid ADHD and CD/ODD: never ALSPAC N: 75; current ALSPAC N:38; previous ALSPAC N: 66

**Table 2. Prediction of dropout (current vs. previous ALSPAC participants)**

	N	Prevalence		Prediction of dropout							
		Current	Previous	Unadjusted				Adjusted (N=4070)			
		%	%	OR	95% CI		p	OR	95% CI		p
Child sex (males)	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

Maximum N for current and previous ALSPAC is 3975 and 1140 respectively. Bonferoni adjusted p (.05/11)=.0045

**Table 3. Simple univariable prediction of disruptive behaviour disorder for the current ALSPAC and previous ALSPAC children (dropouts) using factors assessed during pregnancy**

	N	Current*				Previous†				Interaction§			
		OR	95% CI	p		OR	95% CI	p		OR	95% CI	p	
Child sex (males)	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	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

\*260 cases with disruptive behaviour disorder; 3712 controls; † 72 cases with disruptive behaviour disorder; 1058 controls; § Group (current vs. previous ALSPAC); All children: 332 cases and 4770 controls. Bonferoni adjusted p (.05/11)=.0045.

Fig 1: Description of ALSPAC sample – flow chart

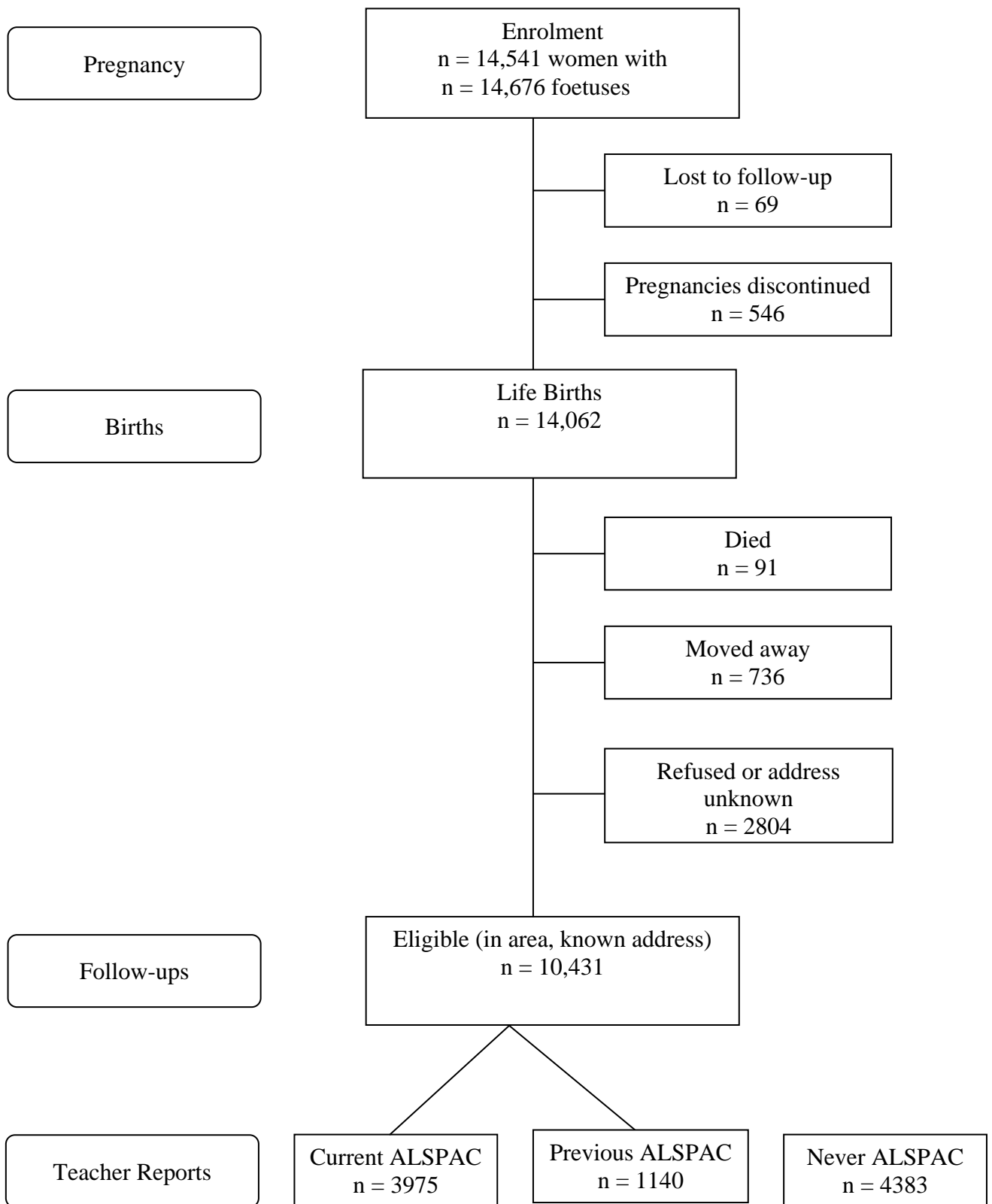
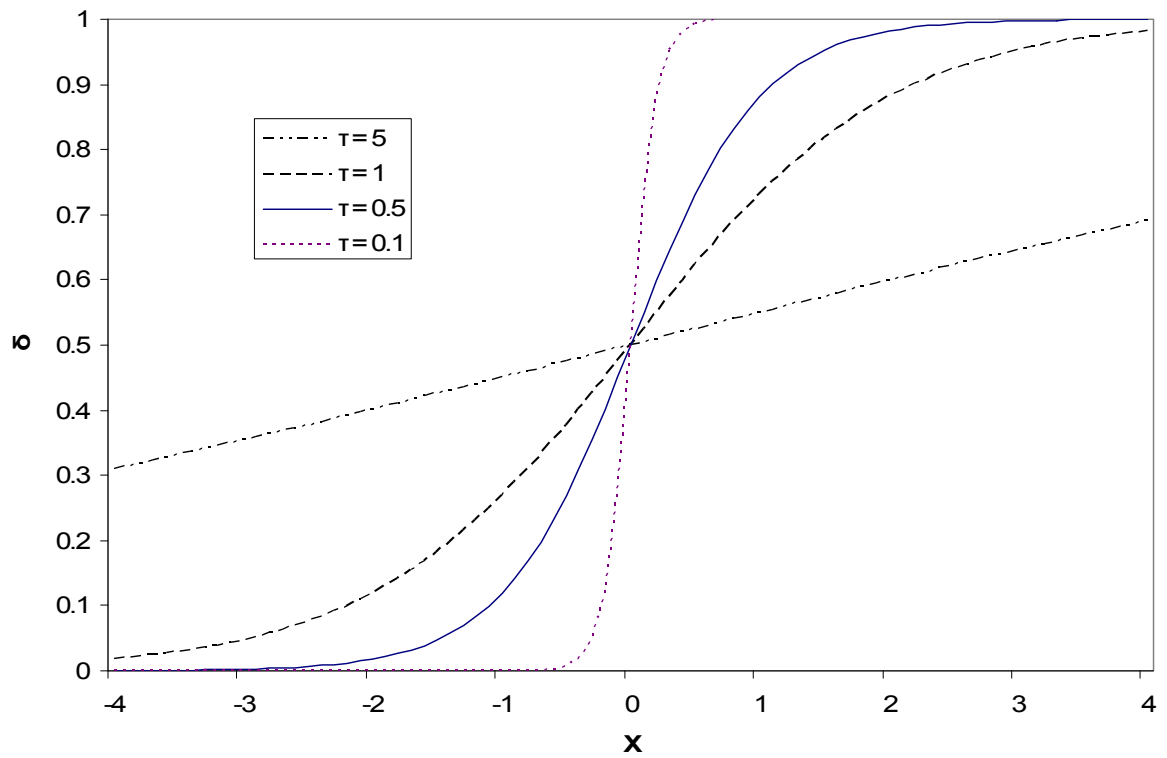
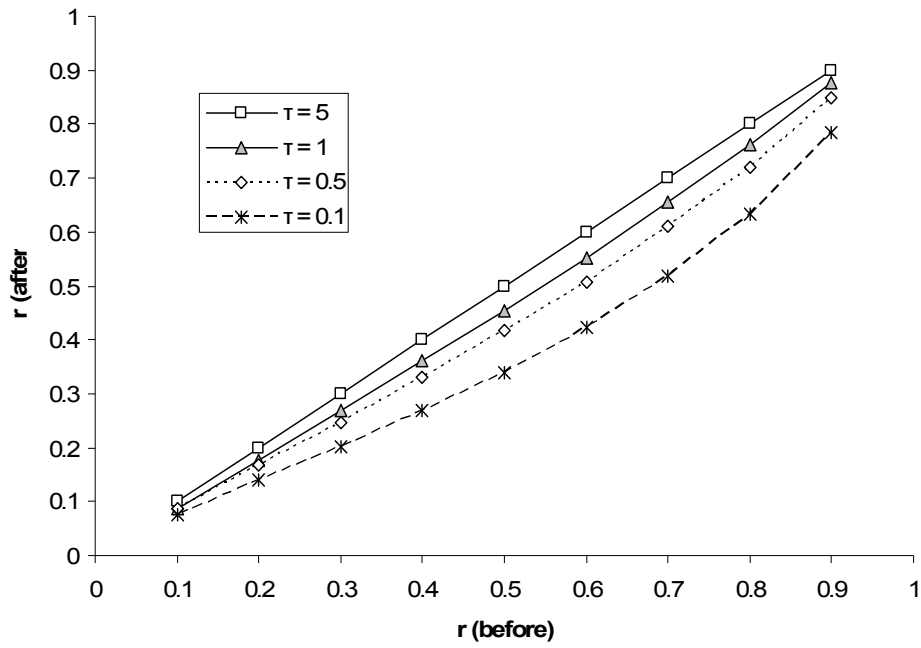


Figure 2 Dropout probability ( $\delta$ ) as a function of  $X$ , for different values of  $\tau$ .

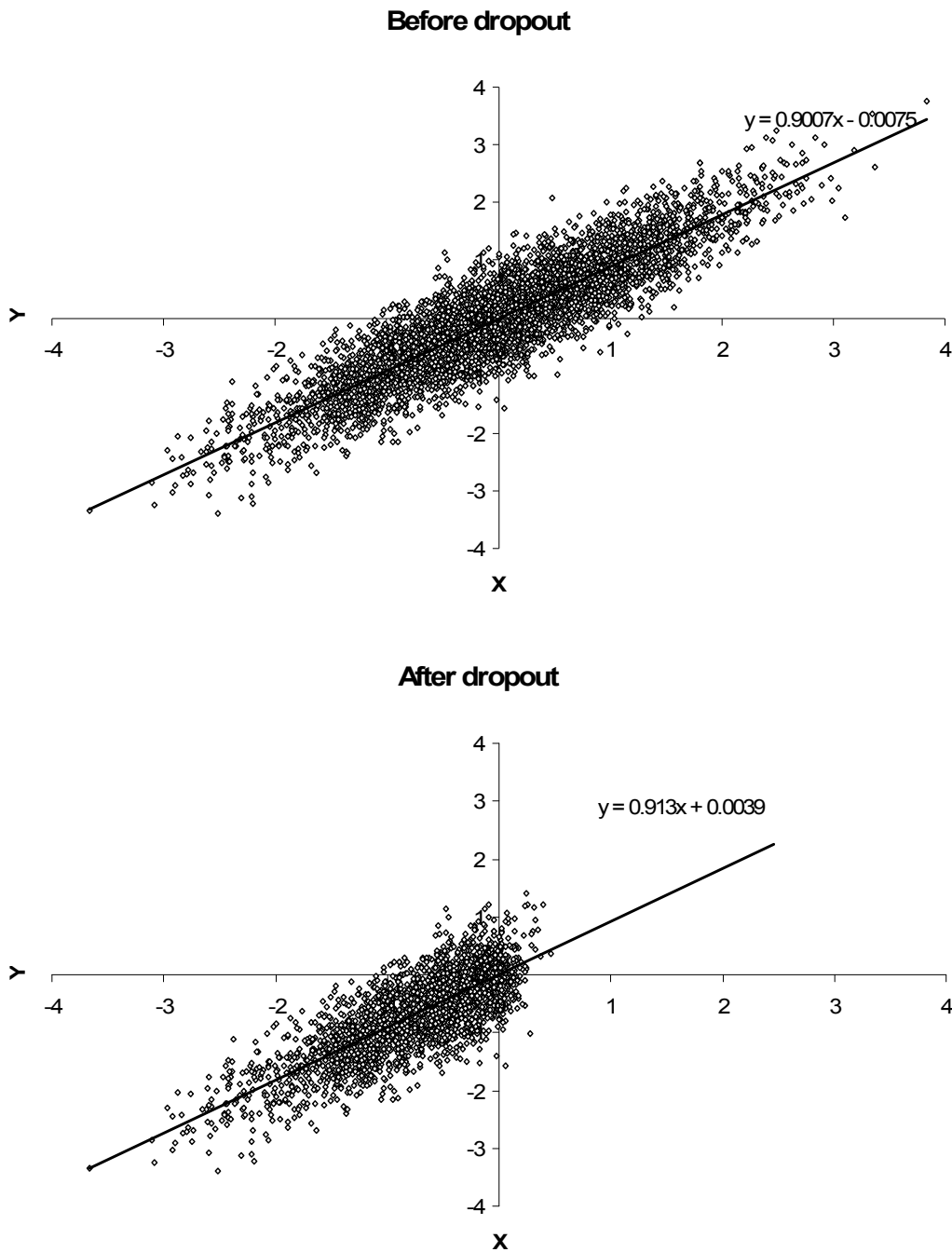


**Figure 3. Correlation between predictor X and criterion Y before and after dropout, as a function of  $\tau$ .**



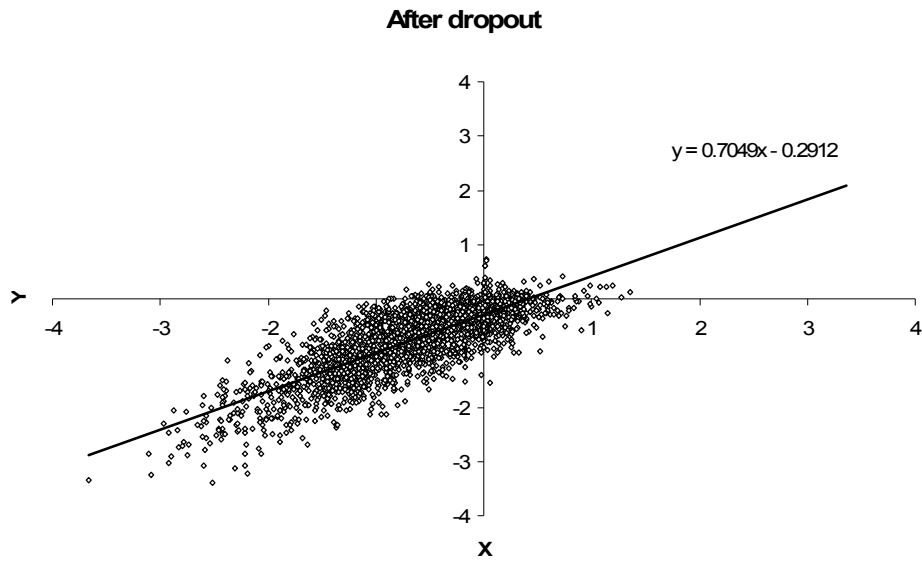
**Figure 4. Simulated effect of selective dropout according to the predictor variable X on least-squares linear regression model.**

**X = predictor; Y = criterion.**

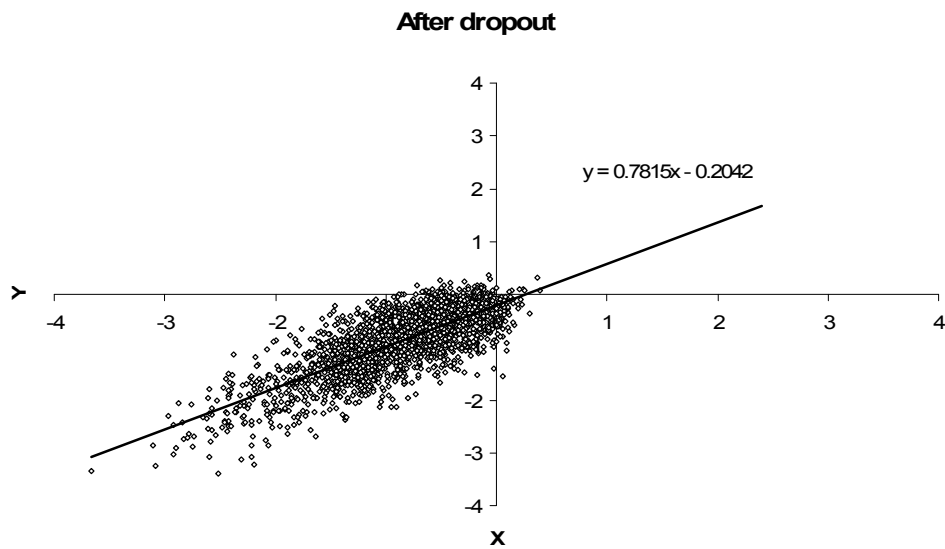


**Figure 5a. Simulated effect of selective dropout according to the criterion variable y on least-squares linear regression model.**

**X = predictor; Y = criterion**



**Figure 5b. Simulated effect of selective dropout according to the predictor variable x and criterion variable y on least-squares linear regression model.**





### **Clinical Implications**

- Cohort or longitudinal studies provide the strongest evidence that a potential risk factor is implicated in the aetiology of an outcome, due to the separation of exposure to the risk factor and the outcome in time. Conventional thinking states that if people with characteristics related to the outcome of interest tend to drop out (selection bias) the findings of the study may be erroneous. This study of disruptive behaviour disorders diagnosed from teacher data and subsequent simulations suggest that systematic drop out does not seriously bias the findings of cohort studies in which significant attrition according to predictor variables has occurred.

### **Limitations**

- Our findings relate to one outcome in one study, and need replication by other groups.
- The response rate among the drop outs and current ALSPAC children was relatively low, and there may have been undetected selection biases in our sample.

### **Contribution of authors.**

RG conceived the idea of comparing the teacher-based diagnoses across the three groups of children to study selection bias. All authors were involved in subsequent discussions in which the basic idea was expanded and developed into its current form.

AW, CS and MS selected, cleaned and collated the data. CS collated and summarised data on response rates from schools.

AW, MS, CS and DW performed the analysis.

KL conducted and summarized the simulations.

DW and TF lead the writing of the manuscript, but all authors collaborated in the writing of the finished paper.