BMJ Open

Impact of participant attrition on child injury outcome estimates in an Australian longitudinal birth cohort

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015584
Article Type:	Research
Date Submitted by the Author:	17-Dec-2016
Complete List of Authors:	Cameron, Cate; Griffith University, Menzies Health Institute Osborne, Jodie; Griffith University Spinks, Annelies; CSIRO Davey, TM; The University of Queensland Sipe, Neil; University of Queensland, School of Geography, Planning and Environmental Management McClure, Rod; Harvard School of Public Health, Harvard Injury Control Research Center
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Research methods
Keywords:	EPIDEMIOLOGY, STATISTICS & RESEARCH METHODS, PUBLIC HEALTH

SCHOLARONE™
Manuscripts

For peer review only

Impact of participant attrition on child injury outcome estimates in an Australian longitudinal birth cohort

Type of Manuscript: Original Article

Authors: Cate M. Cameron,[a] Jodie M. Osborne,[a] Anneliese B. Spinks,[b,c] Tamzyn M. Davey,[d] Neil Sipe,[e] Roderick J. McClure.[f]

Addresses:

[a] Menzies Health Institute Queensland, Griffith University, Meadowbrook, Australia.

[b] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Dutton Park, Australia.

[c] School of Medicine, Griffith University, Meadowbrook, Australia.

[d] The University of Queensland, School of Public Health, Herston, Australia.

[e] School of Geography, Planning and Environmental Management, The University of Queensland, St Lucia, Australia.

[f] Harvard Injury Control Research Center, Harvard School of Public Health, Boston, USA.

Corresponding Author: Dr Cate Cameron, Menzies Health Institute Queensland, Griffith University, University Drive, Meadowbrook, Australia, 4131. Email: cate.cameron@griffith.edu.au, Telephone (07) 338 21289, Facsimile (07) 338 21414.

Word Count: Abstract: 248 words; Text: 2993

Authorship contributions: CMC, JMO, ABS, TMD, NS, RJM were involved in study concept and design. JMO and CMC acquired the data. CMC conducted the data management and analysis. CMC, JMO and RJM were all involved in interpretation of data wrote the first draft of the manuscript. CMC, JMO, ABS, TMD, NS, RJM contributed to the critical revision of the manuscript and approved the final version.

KEY WORDS

Attrition, child injury, epidemiology, longitudinal research, data linkage, methods

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ABSTRACT

Background

Longitudinal research is subject to participant attrition. Systemic differences between retained participants and those lost to attrition potentially bias prevalence of outcomes, as well as exposure-outcome associations. This study examines the impact of attrition on the prevalence of child injury outcomes and the association between sociodemographic factors and child injury.

Methods

Participants were recruited as part of the Environments for Healthy Living (EFHL) birth cohort study. Baseline data was drawn from maternal surveys. Child injury outcome data was extracted from hospital records, 2006-2013. Participant attrition status was assessed up to 2014. Rates of injury related episodes of care were calculated, taking into account exposure time and Poisson regression was performed to estimate exposure-outcome associations.

Results

Of the 2222 participating families, 799 families (36.0%) had complete follow up data. Those with incomplete data included 137 (6.2%) who withdrew, 308 (13.8%) were lost to follow up and 978 families (44.0%) who were partial/non-responders. Families of lower socio-economic status were less likely to have complete follow up data ($p<0.05$). Systematic differences in attrition did not result in differential child injury outcomes or significant differences between the attrition and non-attrition groups in risk factor effect estimates. Participants who withdrew were the only group to demonstrate differences in child injury outcomes.

Conclusion

This research suggests that even with considerable attrition, if the proportion of participants who withdraw is minimal, overall attrition is unlikely to affect the population prevalence estimate of child injury or measures of association between sociodemographic factors and child injury.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Uses a mixed active-passive cohort design from a birth cohort study that employed both data linkage and direct participant follow-up to ascertain comprehensive exposure and outcome data.
- Ability to examine impact of participant attrition on outcome estimates using complete data.
- Relatively small sample size. The adjusted model removed significance even though there was little change in the point estimate and could represent a type 2 error.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

INTRODUCTION

Longitudinal birth cohort studies are invaluable for exploring the impact of early environments (physical, social and economic) on a child's ongoing health and development.[1] One of the greatest threats to the validity of this research design is bias due to participant attrition.[2] Attrition can occur as a result of a participant actively withdrawing from the study, relocating and becoming lost to follow up or not responding to follow up waves.

Research examining attrition in longitudinal studies has identified systemic differences between responders and non-responders across a range of individual, family, social, demographic and health characteristics, with disadvantaged populations generally having higher rates of attrition.[3-8] Systemic characteristic differences have also been found between participant groups according to the reason for attrition.[6, 9, 10] In line with epidemiological principles, it is anticipated that these differences will result in a bias of outcome prevalence and exposure-outcome associations. However, whilst evidence of a link between attrition and prevalence bias has been reported in the literature, there is less evidence of attrition leading to bias in exposure-outcome associations, and the bias effect can vary depending on the nature of the outcomes under investigation.[3-5, 7, 11]

Within the context of longitudinal birth studies, the impact that attrition bias has on child injury outcomes is an important area for further research. To date, the majority of attrition-injury studies have focused on participants recruited after injury has already occurred, or older populations.[12-22] In these studies, attrition has been linked to socio-economic and participant characteristics, causes of injury, injury severity and treatment plans. Extending this research to include longitudinal research on outcomes for young children in a broader population will greatly enhance our understanding of this relationship.

Using data linkage of routinely collected administrative health records with participant survey data, provides researchers new opportunities to assess the impact of attrition bias on participant outcomes in prospective longitudinal studies.[23] Data linkage can provide comprehensive access to outcomes

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

data for the total cohort group, allowing comparisons of outcomes to be made between participants who are active and those who have been lost from 'active' (survey) follow up.

In this manuscript we report the use of a mixed direct participant contact (including surveys) and linked data design in a longitudinal birth cohort study to examine the impact of participant attrition bias on child injury outcome estimates. In particular, the research will i) identify whether there are systematic differences in attrition, and ii) examine the impact of attrition on both the prevalence of child injury outcomes and the association between sociodemographic variables and child injury.

METHODS

Study Design

This cohort study combined survey and linked administrative health data.

Environments for Healthy Living (EFHL): Griffith Birth Cohort Study

The EFHL study is a longitudinal birth cohort which investigates how social, environmental, neighbourhood, family, maternal and individual factors impact a child's health and development.[24] Pregnant women in their third trimester were recruited between 2006 and 2011 at one of three public maternity hospitals in South East Queensland and Northern New South Wales. EFHL has been included on the Australian and New Zealand Clinical Trials Registry (ACTRN12310000931077). The study methodology and baseline characteristics of the sample have been comprehensively described in other papers.[24, 25] The EFHL sample was largely characteristic of births in the region, however, consistent with the public hospital setting from which participants were recruited, there was a higher than national average representation of families with lower incomes, younger maternal age, more overseas born parents and higher proportions of maternal smoking in pregnancy.[25]

Participants

Participants were families enrolled in the EFHL project at one of the two Queensland based public hospitals in the region (Logan or Gold Coast Hospital). This allowed direct comparison to Queensland

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Health child hospital records. Participants were selected from those recruited in the years 2006 to 2010 only, as these cohorts had completed more than one follow up survey. Cases of maternal and child death were excluded.

Ethical considerations

Ethics approval was provided by the participating hospitals and by the Human Research Ethics Committees of Griffith University and the University of Queensland.

Data sharing

No additional data available.

Participant consent

Written, informed maternal consent for data used in this analysis was obtained for (1) direct participant contact including the completion of baseline and follow up surveys; (2) release of hospital perinatal records and (3) for data to be linked to the child's administrative health records including hospital admission and emergency department presentations.

Data Sources and Collection

Baseline Social, Maternal and Household Variables

Baseline information was collected from hospital perinatal records and from a self-administered survey completed by mothers during pregnancy. This baseline survey covered a range of maternal, household and community factors. The child characteristics included in this analysis were gender and age, via length of follow-up. Maternal characteristics included the level of education completed; maternal age at baseline and whether the mother smoked during her pregnancy. Household factors included marital status, whether other children resided in the home, housing status (owned, rented, government/boarding house) and household income level. Household income was divided into quintiles.[25]

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Cohort Attrition Status

Information on participant attrition status was collected from the EFHL participant tracking database. Records were kept on both the participant's overall recruitment status, as well as on their participation in follow up surveys up to May 2014.

Participants were categorised into two groups; (1) those with complete follow-up or (2) those with incomplete follow-up. Participants with complete follow-up included those that had returned all follow-up surveys relevant to their age group and who had not withdrawn during the lifetime of the project.

Participants with incomplete follow-up were categorised into one of three groups. (1) The "withdrawn" group included any participant who had actively asked to be withdrawn from direct participant contact and ongoing survey completion. Consent was retained for the use of existing data and linked administrative data. (2) "Lost to follow-up", included those participants who did not respond to follow-up surveys and who could no longer be contacted through their home address, phone, email or emergency contact addresses. (3) "Partial or Non Responders" included people who were locatable, had not withdrawn, but had not completed one or more follow up surveys beyond the initial baseline survey.

Administrative Health Data

The child's health records were extracted between 2006 and 2013 from the Queensland Emergency Department Information System (EDIS) and from the Queensland Hospital Admitted Patients Data Collection (QHAPDC). The Queensland Department of Health used linkage software, based on deterministic and probabilistic methods, to link demographic, child and maternal information to hospital records. Manual clerical reviews of the data were undertaken by staff when required.

Outcome Measure

Injury related a hospital or Emergency Department (ED) episode of care was the outcome of interest. Multiple admissions, nested admissions and any corresponding hospital or ED presentations relating

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

to the one injury event were identified through dates of presentations, transfer codes combined with diagnostic fields. These records were subsequently collapsed into one episode of care for the injury event.

Injury Classification

Injury presentations were classified using ICD-10-AM, Chapter 19 Injury and Poisonings (S00-T98) and Chapter 20 External Causes (U50-Y98), with late effects from injury excluded. EDIS data contains only a single diagnostic field in which to describe the reason for presentation, whereas QHAPDC contains a primary diagnosis field and multiple other diagnostic fields in which external causes and activity codes may be recorded. Researchers checked text descriptions in the EDIS data to identify any injuries that were not captured in the single diagnosis field. Almost all ED records used the one available diagnostic field to classify the nature of the injury (S00-T98). External cause codes (V01-Y98) were used in less than 5% of ED records. As a result, injury cause codes were only available comprehensively for the inpatient episodes of care. Injury subgroups were defined and matched to the ICD-10-AM chapter subcategories.

Calculation of Person Years

The length of follow up for injury related hospital treatment varied considerably across participants. Individual person-years (PYs) were calculated for each child, based on the time between birth and 31 December 2013 in which he or she was residing in the state of Queensland, alive and eligible for health care.

Analysis

Data cleaning and analyses were undertaken using SAS 9.4 software. The statistical significance of differences between groups was assessed by Pearson's chi-squared test for categorical data. Using the state-wide linked administrative health data, injury related episodes of care were obtained for all cohort participants. Rates of injury related episodes of care were calculated, taking into account PYs exposure time, for each factor and by attrition status (rates/10PYs).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Poisson regression was performed to estimate crude and adjusted rate ratios (RRs) between multiple exposure factors (child, maternal and household characteristics) and the outcome (count of child injury related episodes of care). All factors significant in the univariate analyses associated with injury were included in the model. The log of individual PYs of exposure time was included as an offset. The final model included child gender, maternal age and maternal education. Complete case analysis was employed with two-sided significance set at a level of 5%.

RESULTS

In total, 2222 families who enrolled in the EFHL study from 2006-2010, at the two participating Queensland hospitals, were included in this analysis. The number of children totalled 2245 (including 23 sets of twins). Maternal age ranged from 16 to 48 years with a mean age of 28.9 years (SD± 5.98). At the time of enrolment, 14.1% of the households were sole parent families, 28.2% of maternal participants smoked cigarettes, 34.9% had no other children living in the household and 22.7% of the maternal participants had not completed secondary school.

Attrition Status and Follow-up

Of the 2222 participating families, 799 families (36.0%) had complete follow-up, 137 (6.2%) were withdrawn, 308 (13.8%) were lost to follow-up and 978 families (44.0%) were partial or non-responders (Figure 1).

Insert Figure 1 about here.

Follow-up consisted of a total of 11,908.3 PYs from birth to seven years of age, with a mean 5.3 PYs per child (range 0-7). Automated linkage and manual searching found records in QHAPDC for 97.1% of all EFHL child participants, including their birth record, and records in EDIS for 44.7% of participants.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Baseline Demographic, Household and Child Characteristics by Attrition Status

Baseline demographic and household characteristics differed significantly between participants who had complete follow up compared to those families who had incomplete follow up. Attrition families were more likely to have lower levels of maternal education, lower maternal age, have higher rates of maternal smoking, be single parent households, have more children residing in the home, have lower gross household incomes and live in rental or government boarding housing (all $p<0.0001$). However there were no differences in the proportion of child gender for active families compared to attrition families (Table 1).

Insert Table 1 about here.

These baseline characteristic differences were similar for two of the three attrition groups, lost to follow-up and partial/non-responders. However, those families in the withdrawn group were more similar in baseline characteristics to the families in the complete follow up group, with the exception of maternal age, marital status and gross household income (Table 1). Of note, families in the withdrawn group were more likely to have male children than those families who fully participated ($p<0.01$)

Prevalence Rates of Childhood Injury by Attrition Status

The total cohort had an overall child injury rate of 2.59/10PYs, similar to the child injury rate for those families with complete follow up (2.60/10PYs) and those families who were lost to follow-up or partial/non-responders (Table 2). However, families in the withdrawn group had significantly higher rates of child injury compared to families with complete follow up with an unadjusted RR=1.32 (95% CI 1.03-1.68). While there was little reduction in the effect size after adjusting for co-variates, the difference in child injury rates for families who withdrew compared to families with complete follow up was no longer statistically significant, with an adjusted RR=1.24 (95% CI 0.96-1.58). (Table 2).

Insert Table 2 about here.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Relationship between Baseline Sociodemographic Variables (exposure) and Child Injury Episodes (outcome) by Attrition Status

For the total cohort, rates of child injury increased as the level of maternal education, maternal age and gross household income decreased (Table 3). Rates of child injury were higher in single parent households, families living in government or boarding house accommodation and for families with male children. This distribution was similarly reflected in both the families with complete follow up and those lost to attrition, with no statistically significant differences in the demographically stratified rates of child injury when comparing overall attrition status ($p>0.05$) (Table 3).

Insert Table 3 about here.

While the direction of the demographically stratified child injury rates was similar for all attrition groups, the magnitude of this difference was greater in the withdrawn group for a number of baseline factors. Child injury rates for families that withdrew were higher than families with complete follow up for all baseline factors, but most pronounced with respect to child gender, maternal education, maternal smoking, number of children living in the household and home ownership ($p<0.05$).

DISCUSSION

This study examined the impact of attrition in child injury research on the prevalence of child injury outcomes and the association between sociodemographic variables and injury (exposure-outcome relationship). There were three key findings. First, participants with incomplete follow up, as a group, differed from complete responders across most sociodemographic characteristics and there were also differences in these characteristics across attrition types (within the incomplete follow up group). Second, despite these systematic differences, there were no statistically significant impacts on the overall injury prevalence estimates and risk factor effect estimates, with the exception of the withdrawn group. The withdrawn group differed least, however, from the complete group in sociodemographic characteristics. Third, despite these systemic differences, the direction and nature of the sociodemographic-child injury relationship did not vary across the attrition groups.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The finding that there were systemic differences in the baseline social, maternal, and household characteristics of those with complete follow up versus those lost to attrition is consistent with previous research in this area, with families categorised as having lower socioeconomic status less likely to remain as active participants in the study.[3, 5, 9, 10] Systematic differences in attrition were also evident across attrition groups, as has been demonstrated in previous research.[9, 10] Interestingly, whilst families from the 'lost to follow up' and 'partial/non responder' groups were more likely to be of lower socioeconomic status, our withdrawn participants were more similar in their baseline characteristics to participants with complete follow up. It may be that families of higher sociodemographic status felt more confident to actively withdraw from this research or that there are unmeasured characteristics in these families that lead to higher rates of withdrawal. Further research is warranted given the relatively small sample size of this withdrawn group.

Consistent with previous research,[3, 4, 26] it was anticipated that systemic differences in attrition would result in differences in the prevalence of child injury outcomes across our groups (attrition-child injury), particularly as the relationship between lower sociodemographic factors and child injury has been well documented in the literature.[27-30] However, there were no statistically significant impacts of this differential attrition on the overall injury prevalence estimates and risk factor effect estimates. The only apparent suggestion of difference (albeit with the observed stable effect losing significance with inclusion of multiple variables in the model) was in the withdrawn group, despite this group differing least from the "complete follow up" group for all measures of SES, compared to the other attrition categories. It should be noted though that the withdrawn group was small in absolute terms (i.e. <5% of the total cohort) and had the group been larger, it could have impacted on the overall prevalence of the outcome. This suggests that reasons for attrition may need to be examined in epidemiological research, as a high proportion of withdrawals may indicate the potential for bias.

Importantly, the systemic differences in attrition also had little impact on the direction of the relationship between sociodemographic factors and child injury, with lower socioeconomic status families having higher rates of injury across all the attrition groups. This finding adds to a growing body of evidence indicating that selective attrition does not necessarily impact the relationship

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

between a range of exposure-outcome measures.[3, 4, 10] With follow up rates for cohort studies regularly being reported in the 30-70% range,[31] it is promising that these results suggest exposure-outcome findings from studies with large amounts of attrition may not need to be interpreted with the degree of caution currently applied.

A key strength of this study was the mixed active-passive nature of the cohort design. Prior to the first published results from this study, there had been few birth cohort studies that employed data linkage methodology in combination with active participant contact (survey follow up) to ascertain comprehensive exposure and outcome data. Thus, this is one of the few studies capable of examining with rigour the questions we have addressed in this research. The key limitation is the relatively small sample size. This was particularly evident in relation to our main finding that it was in the group of participants who withdrew from the study that we found a bias. The adjusted model removed significance of the main finding even though there was little change in the point estimate (from 1.32 to 1.24) and could represent a type 2 error. The second limitation relates to the nature of the study sample, the determinant variables measured and the injury outcome focus of the analysis. While the research results could be expected to apply to similar circumstances, it is unknown how validly the study findings should be applied beyond the study group we examined.

The findings of this research provide support for relaxing one of the most challenging expectations of epidemiological studies; that is, complete participant follow up. Even with considerable attrition, when proportions of participants who withdraw from the study are low, attrition bias is unlikely to affect either the population prevalence estimate or measures of association with key study variables.

ACKNOWLEDGMENTS

The research reported in this publication is part of the Environments for Healthy Living (EFHL): Griffith Study of Population Health (Australian and New Zealand Clinical Trials Registry: ACTRN12610000931077). The EFHL project was conceived by Professor Rod McClure, Dr Cate Cameron, Professor Judy Searle and Professor Ronan Lyons. We are thankful for the contributions of the Project Manager, Rani Scott, and the current and past Database Managers. We gratefully

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

acknowledge the administrative staff, research staff, and the hospital antenatal and birth suite midwives of the participating hospitals for their valuable contributions to the study, in addition to the expert advice provided by Research Investigators throughout the project.

FUNDING

This study was funded by an Australian Research Council Discovery Grant [DP110105423]. Core funding to support EFHL is provided by Griffith University. Dr Cameron was supported by a Public Health Fellowship (ID 428254) from the NHMRC.

COMPETING INTERESTS

The authors do not have any competing interests in relation to this research.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

REFERENCES

1. Western M, Haynes M, Baffour B, et al. The Case for a New Australian Birth Cohort Study: Institute for Social Science Research, University of Queensland; 2015.
2. Ahern K, Le Brocque R. Methodological Issues in the Effects of Attrition: Simple Solutions for Social Scientists. *Field Methods*. 2005;17(1):53-69.
3. Wolke D, Waylen A, Samara M, et al. Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British journal of psychiatry : the journal of mental science*. 2009;195(3):249-56.doi:10.1192/bjp.bp.108.053751.
4. Osler M, Kriegbaum M, Christensen U, et al. Loss to follow up did not bias associations between early life factors and adult depression. *J Clin Epidemiol*. 2008;61(9):958-63.doi:10.1016/j.jclinepi.2007.11.005.
5. Greene N, Greenland S, Olsen J, et al. Estimating bias from loss to follow-up in the Danish National Birth Cohort. *Epidemiology*. 2011;22(6):815-22.doi:10.1097/EDE.0b013e31822939fd.
6. Ng SK, Scott R, Scuffham PA. Contactable Non-responders Show Different Characteristics Compared to Lost to Follow-Up Participants: Insights from an Australian Longitudinal Birth Cohort Study. *Matern Child Health J*. 2016.doi:10.1007/s10995-016-1946-8.
7. Howe LD, Tilling K, Galobardes B, et al. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. *Epidemiology*. 2013;24(1):1-9.doi:10.1097/EDE.0b013e31827623b1.
8. Aylward GP, Hatcher RP, Stripp B, et al. Who goes and who stays: subject loss in a multicenter, longitudinal follow-up study. *J Dev Behav Pediatr*. 1985;6(1):3-8.
9. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? *Aust N Z J Public Health*. 2006;30(4):353-61.
10. Powers J, Tavener M, Graves A, et al. Loss to follow-up was used to estimate bias in a longitudinal study: a new approach. *J Clin Epidemiol*. 2015;68(8):870-6.doi:10.1016/j.jclinepi.2015.01.010.
11. Cornish R, Tilling K, Boyd A, et al. Using linkage to electronic primary care records to evaluate recruitment and nonresponse bias in the Avon Longitudinal Study of Parents and Children. *Epidemiology*. 2015;26(4):e41-2.doi:10.1097/EDE.0000000000000288.
12. Jourdan C, Bayen E, Bahrami S, et al. Loss to follow-up and social background in an inception cohort of patients with severe traumatic brain injury: results from the PariS-TBI study. *The Journal of head trauma rehabilitation*. 2016;31(3):E42-E8.
13. Tin Tin S, Woodward A, Ameratunga S. Estimating bias from loss to follow-up in a prospective cohort study of bicycle crash injuries. *Inj Prev*. 2014;20(5):322-9.doi:10.1136/injuryprev-2013-040997.
14. Holavanahalli RK, Lezotte DC, Hayes MP, et al. Profile of patients lost to follow-up in the Burn Injury Rehabilitation Model Systems' longitudinal database. *J Burn Care Res*. 2006;27(5):703-12.doi:10.1097/01.BCR.0000238085.87863.81.
15. Blaha RZ, Arnett AB, Kirkwood MW, et al. Factors influencing attrition in a multisite, randomized, clinical trial following traumatic brain injury in adolescence. *J Head Trauma Rehabil*. 2015;30(3):E33-40.doi:10.1097/HTR.0000000000000059.
16. Corrigan JD, Harrison-Felix C, Bogner J, et al. Systematic bias in traumatic brain injury outcome studies because of loss to follow-up. *Arch Phys Med Rehabil*. 2003;84(2):153-60.doi:10.1053/apmr.2003.50093.
17. Gabbe BJ, Cleland H, Watterson DM, et al. Long term outcomes data for the Burns Registry of Australia and New Zealand: Is it feasible? *Burns : journal of the International Society for Burn Injuries*. 2015;41(8):1732-40.doi:10.1016/j.burns.2015.09.005.
18. Wade SL, Gerry Taylor H, Yeates KO, et al. Long-term parental and family adaptation following pediatric brain injury. *Journal of pediatric psychology*. 2006;31(10):1072-83.doi:10.1093/jpepsy/jsj077.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

19. Krellman JW, Kolakowsky-Hayner SA, Spielman L, et al. Predictors of follow-up completeness in longitudinal research on traumatic brain injury: findings from the National Institute on Disability and Rehabilitation Research traumatic brain injury model systems program. *Arch Phys Med Rehabil*. 2014;95(4):633-41.doi:10.1016/j.apmr.2013.10.016.
20. de Groot S, Haisma JA, Post MW, et al. Investigation of bias due to loss of participants in a Dutch multicentre prospective spinal cord injury cohort study. *J Rehabil Med*. 2009;41(5):382-9.doi:10.2340/16501977-0346.
21. Neuner B, Fleming M, Born R, et al. Predictors of loss to follow-up in young patients with minor trauma after screening and written intervention for alcohol in an urban emergency department. *Journal of studies on alcohol and drugs*. 2007;68(1):133-40.
22. Lin C-YA, DeMatteo CA. The Challenges of Loss to Follow-up in Longitudinal Pediatric Acquired Brain Injury Research: One Research Team's Experiences. 2013.
23. Calderwood L, Lessof C. Enhancing longitudinal surveys by linking to administrative data. *Methodology of longitudinal surveys*. 2009:55-72.
24. Cameron CM, Scuffham PA, Spinks A, et al. Environments for Healthy Living (EFHL) Griffith birth cohort study: background and methods. *Matern Child Health J*. 2012a;16(9):1896-905.doi:10.1007/s10995-011-0940-4.
25. Cameron CM, Scuffham PA, Shibl R, et al. Environments For Healthy Living (EFHL) Griffith birth cohort study: characteristics of sample and profile of antenatal exposures. *BMC public health*. 2012b;12:1080.doi:10.1186/1471-2458-12-1080.
26. Mars B, Cornish R, Heron J, et al. Using Data Linkage to Investigate Inconsistent Reporting of Self-Harm and Questionnaire Non-Response. *Archives of suicide research : official journal of the International Academy for Suicide Research*. 2016;20(2):113-41.doi:10.1080/13811118.2015.1033121.
27. Hong J, Lee B, Ha EH, et al. Parental socioeconomic status and unintentional injury deaths in early childhood: consideration of injury mechanisms, age at death, and gender. *Accident; analysis and prevention*. 2010;42(1):313-9.doi:10.1016/j.aap.2009.08.010.
28. Kim M-H, Subramanian SV, Kawachi I, et al. Association between childhood fatal injuries and socioeconomic position at individual and area levels: a multilevel study. *J Epidemiol Community Health*. 2007;61:135-40.doi:doi:10.1136/jech.2006.047738.
29. Laursen B, Nielsen JW. Influence of sociodemographic factors on the risk of unintentional childhood home injuries. *European journal of public health*. 2008;18(4):366-70.doi:10.1093/eurpub/ckn034.
30. Myhre MC, Thoresen S, Grogaard JB, et al. Familial factors and child characteristics as predictors of injuries in toddlers: a prospective cohort study. *BMJ open*. 2012;2(2):e000740.doi:10.1136/bmjopen-2011-000740.
31. Gustavson K, von Soest T, Karevold E, et al. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC public health*. 2012;12:918.doi:10.1186/1471-2458-12-918.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

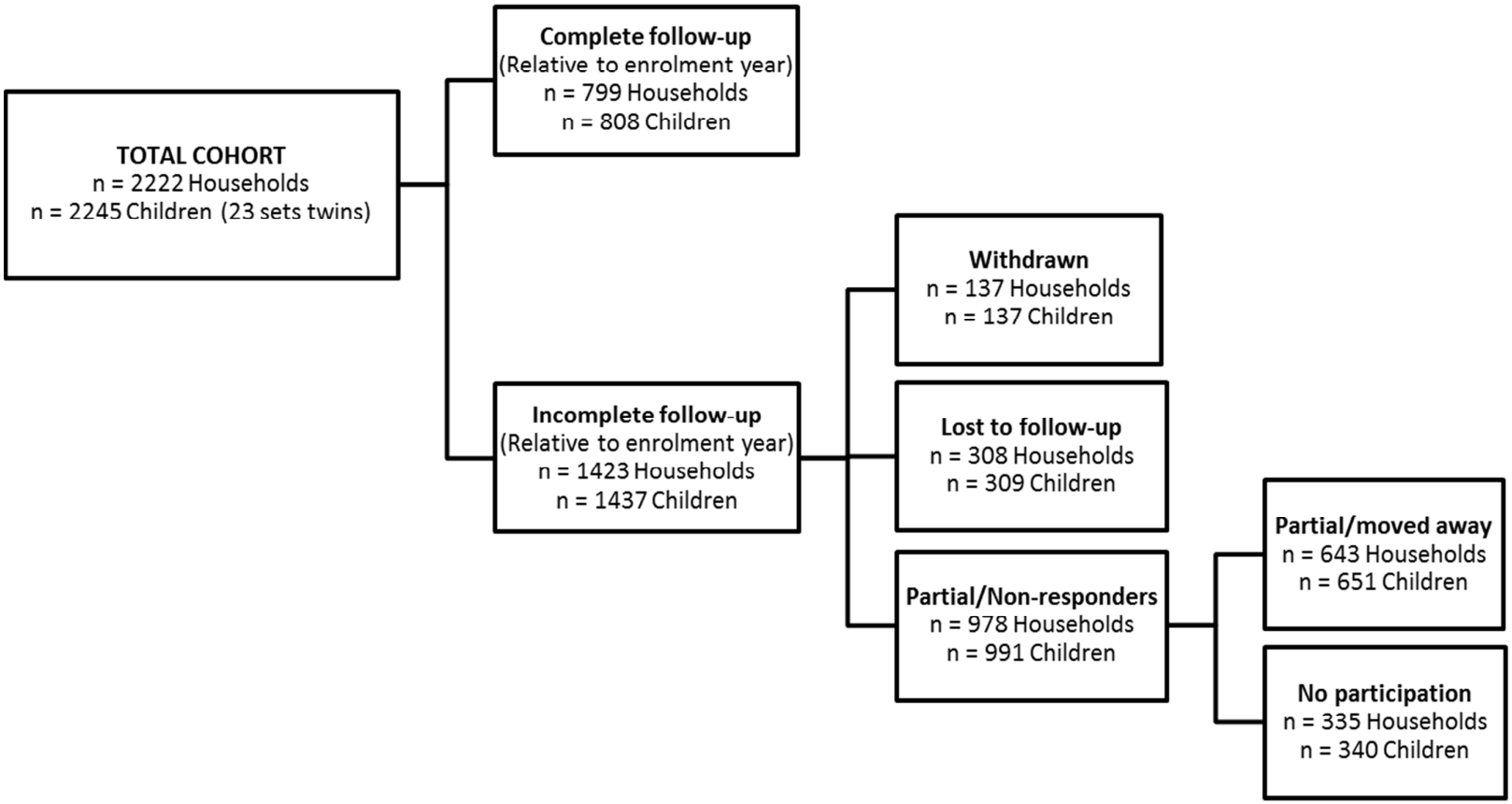

Figure 1 Flow diagram of EFHL cohort attrition status, Queensland participants 2006-2010 only (N=2222 Households and N=2245 Children).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1 Distribution of baseline maternal, household and child characteristics, stratified by follow-up status and types of attrition.

	Follow-up			Attrition by type [a]		
	Complete	Incomplete		Withdrawn	Lost to Follow-up	Partial/Non Responders
	N (%)	N (%)	p-value	N (%)	N (%)	N (%)
MATERNAL CHARACTERISTICS	**799**	**1423**		**137**	**308**	**978**
Maternal level of education *Missing=12*						
Tertiary degree	192 (24.1)	178 (12.6)	***	24 (17.5)	34 (11.0) ***	120 (12.4) ***
Secondary school or trade	490 (61.6)	849 (60.0)		89 (65.0)	182 (59.3)	578 (59.6)
Not complete school	114 (14.3)	387 (27.4)		24 (17.5)	91 (29.6)	272 (28.0)
Maternal age *Missing=3*						
35+ years	201 (25.2)	228 (16.1)	***	17 (12.4) ***	53 (17.2) ***	158 (16.2) ***
30-34 years	253 (31.7)	308 (21.7)		37 (27.0)	56 (18.2)	215 (22.0)
25-29 years	228 (28.6)	416 (29.3)		48 (35.0)	83 (26.9)	285 (29.2)
<25 years	116 (14.5)	469 (33.0)		35 (25.6)	116 (37.7)	318 (32.6)
Smoking during pregnancy *Missing=9*						
No	640 (80.1)	949 (67.1)	***	102 (74.5)	187 (60.9) ***	660 (68.0) ***
Yes	159 (19.9)	465 (32.9)		35 (25.5)	120 (39.1)	310 (32.0)
HOUSEHOLD CHARACTERISTICS	**799**	**1423**		**137**	**308**	**978**
Marital status *Missing=18*						
Two parent family	732 (92.1)	1161 (82.4)	***	106 (77.4) ***	247 (81.2) ***	808 (83.5) ***
Sole parent family	63 (7.9)	248 (17.6)		31 (22.6)	57 (18.8)	160 (16.5)
Other children in household *Missing=14*						
None	299 (37.6)	471 (33.4)	***	53 (39.0)	97 (31.8) ***	321 (33.0) ***
1-2 children	431 (54.2)	731 (51.7)		70 (51.5)	154 (50.5)	507 (52.2)
3+ children	65 (8.2)	211 (14.9)		13 (9.5)	54 (17.7)	144 (14.8)
Household income *Missing=354*						
Highest quintile	170 (24.0)	175 (15.1)	***	17 (15.2) *	30 (12.1) ***	128 (16.0)***
4th quintile	154 (21.7)	206 (17.8)		31 (27.7)	31 (12.4)	144 (18.1)
3rd quintile	164 (23.1)	211 (18.2)		21 (18.7)	42 (16.9)	148 (18.6)
2nd quintile	129 (18.2)	259 (22.3)		19 (17.0)	65 (26.1)	175 (21.9)
Lowest quintile	92 (13.0)	308 (26.6)		24 (21.4)	81 (32.5)	203 (25.4)
Home ownership *Missing=20*						
Own/mortgage	427 (53.6)	437 (31.1)	***	62 (45.3)	49 (16.2) ***	326 (33.7) ***
Rent	297 (37.2)	740 (52.7)		57 (41.6)	203 (67.2)	480 (49.7)
Government/Boarding house	73 (9.2)	228 (16.2)		18 (13.1)	50 (16.6)	160 (16.6)
CHILD CHARACTERISTICS	**808**	**1437**		**137**	**309**	**991**
Gender *Missing = 49*						
Female	400 (50.1)	670 (48.0)	ns	51 (37.5) **	151 (49.7)	468 (48.9)
Male	399 (49.9)	727 (52.0)		85 (62.5)	153 (50.3)	489 (51.1)

Maternal and Household N=2222; Children N=2245 (includes twins)

[a] Each Drop-out participant group compared to the complete participant group as the reference, using chi-squared test.

*p<0.05 ** p<0.01 ***p <0.0001 ns= not significant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2 Number of injuries, injury rates per 10 person-years (PYs) and Injury Rate Ratios (RRs) with 95% Confidence Intervals (CIs) for the child participants, stratified by follow-up status and types of attrition.

Child Participants	N (%)	Sum Injuries	Injury Rates /10 PYs	Unadjusted RR (95% CI) [a]	Adjusted RR (95% CI) [a,b]
Total cohort	2245 (100.0)	941	2.59	0.99 (0.88-1.13)	0.97 (0.85-1.10)
Complete follow-up	808 (36.0)	340	2.60	Reference	Reference
Incomplete follow-up	1437 (64.0)	601	2.58	0.99 (0.87-1.13)	0.95 (0.82-1.09)
Withdrawn	137 (6.1)	80	3.48	**1.32 (1.03-1.68) ***	1.24 (0.96-1.58)
Lost to follow-up	309 (13.8)	128	2.38	0.89 (0.72-1.09)	0.81 (0.65-1.01)
Partial/Non-responders	991 (44.1)	393	2.51	0.98 (0.85-1.13)	0.94 (0.81-1.10)

[a] Poisson regression, offset = log of length of follow-up time of child.

[b] Adjusted for maternal age, maternal education, maternal smoking status, number of children in household, family marital status, household income, housing status and child gender in two step model. Final model included maternal age, maternal education level and child gender

*p<0.05

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3 Child injury rates per 10 person-years (PYs) for the total cohort and stratified by follow-up status and types of attrition for baseline maternal, household and child characteristics.

	Total Cohort	Follow-up			Attrition by type [a]		
		Complete	Incomplete		Withdrawn	Lost to Follow-up	Partial/Non Responders
	/10 PYs	/10 PYs	/10 PYs	p-value	/10 PYs	/10 PYs	/10 PYs
MATERNAL CHARACTERISTICS							
Maternal level of education *Missing=12*							
Tertiary degree	2.03	2.12	1.92	ns	3.06 *	1.72	1.74
Secondary school or trade	2.70	2.84	2.62		3.66	2.28	2.57
Not complete school	2.73	2.43	2.83		3.23	2.87	2.77
Maternal age *Missing=3*							
35+ years	2.02	1.83	2.18	ns	2.83	2.76	1.91
30-34 years	2.67	2.86	2.52		3.93	1.50	2.55
25-29 years	2.27	2.35	2.23		2.93	1.69	2.27
<25 years	3.27	3.76	3.14		4.12	3.13	3.04
Smoking during pregnancy *Missing=9*							
No	2.49	2.48	2.50	ns	3.25 *	2.55	2.36
Yes	2.84	3.10	2.75		4.17	2.15	2.84
HOUSEHOLD CHARACTERISTICS							
Marital status *Missing=18*							
Two parent family	2.45	2.53	2.41	ns	3.36	2.33	2.30
Sole parent family	3.44	3.47	3.43		3.90	2.66	3.64
Other children in household *Missing=14*							
None	2.70	2.84	2.61	ns	3.09 *	2.66	2.51
1-2 children	2.52	2.51	2.53		3.50	2.30	2.46
3+ children	2.52	1.99	2.68		4.21	2.17	2.73
Household income [b] *Missing=354*							
Highest quintile	2.03	2.42	1.67	ns	2.04	1.52	1.65
4th quintile	2.55	2.93	2.27		2.79	1.95	2.24
3rd quintile	2.45	2.33	2.55		3.57	2.92	2.27
2nd quintile	2.54	2.32	2.65		3.74	2.32	2.66
Lowest quintile	2.92	3.53	2.74		3.15	2.36	2.85
Home ownership *Missing=20*							
Own/mortgage	2.41	2.53	2.29	ns	3.68 *	1.47	2.16
Rent	2.46	2.48	2.46		2.54	2.50	2.43
Government/Boarding house	3.55	3.39	3.60		5.85	2.93	3.57
CHILD CHARACTERISTICS							
Gender *Missing = 49*							
Female	2.18	2.21	2.16	ns	3.13 *	1.89	2.15
Male	3.00	2.87	3.07		3.74	2.93	2.99

Children N=2245 (includes twins)

[a] Each Drop-out participant group compared to the complete participant group as the reference, using univariate Poisson regression with time offset.

[b] Gross annual household incomes standardised in AUD$2010 values

*p<0.05 ** p<0.01 ***p <0.0001 ns=not significant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	PAGE
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	Na
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	**Paper about bias**
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	8-9
		(*c*) Explain how missing data were addressed	9
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	**Paper about bias from loss to follow-up**
		(*e*) Describe any sensitivity analyses	

Continued on next page

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Figure 1
		(b) Give reasons for non-participation at each stage	Figure 1
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	9-10
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	Table 3
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2, page 9
		(*b*) Report category boundaries when continuous variables were categorized	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11
Discussion			
Key results	18	Summarise key results with reference to study objectives	11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

Impact of participant attrition on child injury outcome estimates: a longitudinal birth cohort study in Australia

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015584.R1
Article Type:	Research
Date Submitted by the Author:	23-Mar-2017
Complete List of Authors:	Cameron, Cate; Griffith University, Menzies Health Institute Osborne, Jodie; Griffith University Spinks, Annelies; CSIRO Davey, TM; The University of Queensland Sipe, Neil; University of Queensland, School of Geography, Planning and Environmental Management McClure, Rod; Harvard School of Public Health, Harvard Injury Control Research Center
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Research methods
Keywords:	EPIDEMIOLOGY, STATISTICS & RESEARCH METHODS, PUBLIC HEALTH

SCHOLARONE™
Manuscripts

For peer review only

Impact of participant attrition on child injury outcome estimates: a longitudinal birth cohort study in Australia

Type of Manuscript: Original Article

Authors: Cate M. Cameron,[a] Jodie M. Osborne,[a] Anneliese B. Spinks,[b,c] Tamzyn M. Davey,[d] Neil Sipe,[e] Roderick J. McClure.[f]

Addresses:

[a] Menzies Health Institute Queensland, Griffith University, Meadowbrook, Australia.

[b] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Dutton Park, Australia.

[c] School of Medicine, Griffith University, Meadowbrook, Australia.

[d] The University of Queensland, School of Public Health, Herston, Australia.

[e] School of Geography, Planning and Environmental Management, The University of Queensland, St Lucia, Australia.

[f] Harvard Injury Control Research Center, Harvard School of Public Health, Boston, USA.

Corresponding Author: Dr Cate Cameron, Menzies Health Institute Queensland, Griffith University, University Drive, Meadowbrook, Australia, 4131. Email: cate.cameron@griffith.edu.au, Telephone (07) 338 21289, Facsimile (07) 338 21414.

Word Count: Abstract: 248 words; Text: 2993

Authorship contributions: CMC, JMO, ABS, TMD, NS, RJM were involved in study concept and design. JMO and CMC acquired the data. CMC conducted the data management and analysis. CMC, JMO and RJM were all involved in interpretation of data wrote the first draft of the manuscript. CMC, JMO, ABS, TMD, NS, RJM contributed to the critical revision of the manuscript and approved the final version.

KEY WORDS

Attrition, child injury, epidemiology, longitudinal research, data linkage, methods

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

ABSTRACT

Background

Longitudinal research is subject to participant attrition. Systemic differences between retained participants and those lost to attrition potentially bias prevalence of outcomes, as well as exposure-outcome associations. This study examines the impact of attrition on the prevalence of child injury outcomes and the association between sociodemographic factors and child injury.

Methods

Participants were recruited as part of the Environments for Healthy Living (EFHL) birth cohort study. Baseline data were drawn from maternal surveys. Child injury outcome data were extracted from hospital records, 2006-2013. Participant attrition status was assessed up to 2014. Rates of injury related episodes of care were calculated, taking into account exposure time and Poisson regression was performed to estimate exposure-outcome associations.

Results

Of the 2222 participating families, 799 families (36.0%) had complete follow up data. Those with incomplete data included 137 (6.2%) who withdrew, 308 (13.8%) were lost to follow up and 978 families (44.0%) who were partial/non-responders. Families of lower socio-economic status were less likely to have complete follow up data ($p<0.05$). Systematic differences in attrition did not result in differential child injury outcomes or significant differences between the attrition and non-attrition groups in risk factor effect estimates. Participants who withdrew were the only group to demonstrate differences in child injury outcomes.

Conclusion

This research suggests that even with considerable attrition, if the proportion of participants who withdraw is minimal, overall attrition is unlikely to affect the population prevalence estimate of child injury or measures of association between sociodemographic factors and child injury.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Uses a mixed active-passive cohort design from a birth cohort study that employed both data linkage and direct participant follow-up to ascertain comprehensive exposure and outcome data.
- Ability to examine impact of participant attrition on outcome estimates using complete data.
- Relatively small sample size. The adjusted model removed significance even though there was little change in the point estimate and could represent a type 2 error.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

INTRODUCTION

Longitudinal birth cohort studies are invaluable for exploring the impact of early environments (physical, social and economic) on a child's ongoing health and development.[1] One of the greatest threats to the validity of this research design is bias due to participant attrition.[2] Attrition can occur as a result of a participant actively withdrawing from the study, relocating and becoming lost to follow up or not responding to follow up waves.

Research examining attrition in longitudinal studies has identified systemic differences between responders and non-responders across a range of individual, family, social, demographic and health characteristics, with disadvantaged populations generally having higher rates of attrition.[3-8] Systemic characteristic differences have also been found between participant groups according to the reason for attrition.[6 9 10] In line with epidemiological principles, it is anticipated that these differences will result in a bias of outcome prevalence and exposure-outcome associations. However, whilst evidence of a link between attrition and prevalence bias has been reported in the literature, there is less evidence of attrition leading to bias in exposure-outcome associations, and the bias effect can vary depending on the nature of the outcomes under investigation.[3-5 7 11]

Within the context of longitudinal birth studies, the impact that attrition bias has on child injury outcomes is an important area for further research. To date, the majority of attrition-injury studies have focused on participants recruited after injury has already occurred, or older populations.[12-22] In these studies, attrition has been linked to socio-economic and participant characteristics, causes of injury, injury severity and treatment plans. Extending this research to include longitudinal research on outcomes for young children in a broader population will greatly enhance our understanding of this relationship.

Using data linkage of routinely collected administrative health records with participant survey data, provides researchers new opportunities to assess the impact of attrition bias on participant outcomes in prospective longitudinal studies.[23] Data linkage can provide comprehensive access to outcomes

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

data for the total cohort group, allowing comparisons of outcomes to be made between participants who are active and those who have been lost from 'active' (survey) follow up.

In this manuscript we report the use of a mixed direct participant contact (including surveys) and linked data design in a longitudinal birth cohort study to examine the impact of participant attrition bias on child injury outcome estimates. In particular, the research will i) identify whether there are systematic differences in attrition, and ii) examine the impact of attrition on both the prevalence of child injury outcomes and the association between sociodemographic variables and child injury.

METHODS

Study Design

This cohort study combined survey and linked administrative health data.

Environments for Healthy Living (EFHL): Griffith Birth Cohort Study

The EFHL study is a longitudinal birth cohort which investigates how social, environmental, neighbourhood, family, maternal and individual factors impact a child's health and development.[24] Pregnant women in their third trimester were recruited between 2006 and 2011 at one of three public maternity hospitals in South East Queensland and Northern New South Wales. Participants completed a questionnaire at baseline and then follow-up surveys were sent 1, 3 and 5 years after the birth of the child. In between these scheduled contacts, regular newsletters were sent and additional sub-study project contacts occurred. Returned mail or contact difficulties triggered the use of alternative contact mechanisms supplied at enrolment and updated at follow-ups, including a relative or friends contact details, email and Facebook. [24 25]The EFHL sample was largely characteristic of births in the region, however, consistent with the public hospital setting from which participants were recruited, there was a higher than national average representation of families with lower incomes, younger maternal age, more overseas born parents and higher proportions of maternal smoking in pregnancy.[25] The study methodology and baseline characteristics of the sample have been comprehensively described in other

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

papers.[24 25] EFHL has been included on the Australian and New Zealand Clinical Trials Registry (ACTRN12310000931077).

Participants

Participants were families enrolled in the EFHL project at one of the two Queensland based public hospitals in the region (Logan or Gold Coast Hospital). This allowed direct comparison to Queensland Health child hospital records. Participants were selected from those recruited in the years 2006 to 2010 only, as these cohorts had completed more than one follow up survey. Cases of maternal and child death were excluded from the analysis as these participants had died (which is a known outcome) and were not lost to follow up.

Ethical considerations

Ethics approval was provided by the participating hospitals and by the Human Research Ethics Committees of Griffith University and the University of Queensland.

Data sharing

No additional data available.

Participant consent

Written, informed maternal consent for data used in this analysis was obtained for (1) direct participant contact including the completion of baseline and follow up surveys; (2) release of hospital perinatal records and (3) for data to be linked to the child's administrative health records including hospital admission and emergency department presentations.

Data Sources and Collection

Baseline Social, Maternal and Household Variables

Baseline information was collected from hospital perinatal records and from a self-administered survey completed by mothers during pregnancy. This baseline survey covered a range of maternal,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

household and community factors. The child characteristics included in this analysis were gender and age, via length of follow-up. Maternal characteristics included the level of education completed; maternal age at baseline and whether the mother smoked during her pregnancy. Household factors included marital status, whether other children resided in the home, housing status (owned, rented, government/boarding house) and household income level. Household income was divided into quintiles.[25]

Cohort Attrition Status

Information on participant attrition status was collected from the EFHL participant tracking database. Records were kept on both the participant's overall recruitment status, as well as on their participation in follow up surveys up to May 2014.

Participants were categorised into two groups; (1) those with complete follow-up or (2) those with incomplete follow-up. Participants with complete follow-up included those that had returned all follow-up surveys relevant to their age group and who had not withdrawn during the lifetime of the project.

Participants with incomplete follow-up were categorised into one of three groups. (1) The "withdrawn" group included any participant who had actively asked to be withdrawn from direct participant contact and ongoing survey completion. Consent was retained for the use of existing data and linked administrative data. (2) "Lost to follow-up", included those participants who did not respond to follow-up surveys and who could no longer be contacted through their home address, phone, email or emergency contact addresses. (3) "Partial or Non Responders" included people who were locatable, had not withdrawn, but had not completed one or more follow up surveys beyond the initial baseline survey.

Administrative Health Data

The child's health records were extracted between 2006 and 2013 from the Queensland Emergency Department Information System (EDIS) and from the Queensland Hospital Admitted Patients Data

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Collection (QHAPDC). The matching procedure was undertaken by the Queensland Department of Health (the custodians of the data) using linkage software, based on deterministic and probabilistic methods, to link demographic, child and maternal information to hospital records. Deterministic linkage involves the linking of data sets through comparing fields such as name, year of birth and street name with the requirement that the records agree on all characters. Probabilistic linkage involves the use of statistical models and algorithms to estimate the probability of data from different data sets having commonality (e.g. the same person/event). Clerical reviews of the data were undertaken to manually inspect uncertain matches in probabilistic linkage.[26]

Outcome Measure

Injury related to a hospital or Emergency Department (ED) episode of care was the outcome of interest. Multiple admissions, nested admissions and any corresponding hospital or ED presentations relating to the one injury event were identified through dates of presentations, transfer codes combined with diagnostic fields. These records were subsequently collapsed into one episode of care for the injury event.

Injury Classification

Injury presentations were classified using ICD-10-AM, Chapter 19 Injury and Poisonings (S00-T98) and Chapter 20 External Causes (U50-Y98), with late effects from injury excluded. EDIS data contains only a single diagnostic field in which to describe the reason for presentation, whereas QHAPDC contains a primary diagnosis field and multiple other diagnostic fields in which external causes and activity codes may be recorded. Researchers checked text descriptions in the EDIS data to identify any injuries that were not captured in the single diagnosis field. Almost all ED records used the one available diagnostic field to classify the nature of the injury (S00-T98). External cause codes (V01-Y98) were used in less than 5% of ED records. As a result, injury cause codes were only available comprehensively for the inpatient episodes of care. Injury subgroups were defined and matched to the ICD-10-AM chapter subcategories.

Calculation of Person Years

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

The ages of the children and length of follow up for injury related hospital treatment varied considerably across participants due to the 5 recruitment waves. As such, individual person-years (PYs) of exposure time were calculated for each child, based on the time between birth and 31 December 2013 in which he or she was residing in the state of Queensland, alive and eligible for health care.

Analysis

Data cleaning and analyses were undertaken using SAS 9.4 software. The statistical significance of differences between groups was assessed by Pearson's chi-squared test for categorical data. Using the state-wide linked administrative health data, injury related episodes of care were obtained for all cohort participants. Rates of injury related episodes of care were calculated, taking into account PYs exposure time, for each factor and by attrition status (rates/10PYs).

Poisson regression was performed to estimate crude and adjusted rate ratios (RRs) between multiple exposure factors (child, maternal and household characteristics) and the outcome (count of child injury related episodes of care). All factors significant in the univariate analyses associated with injury were included in the model. The log of individual PYs of exposure time was included as an offset. The final model included child gender, maternal age and maternal education. Complete case analysis was employed with two-sided significance set at a level of 5%.

RESULTS

In total, 2222 families who enrolled in the EFHL study from 2006-2010, at the two participating Queensland hospitals, were included in this analysis. The number of children totalled 2245 (including 23 sets of twins). Maternal age ranged from 16 to 48 years with a mean age of 28.9 years (SD± 5.98). At the time of enrolment, 14.1% of the households were sole parent families, 28.2% of maternal participants smoked cigarettes, 34.9% had no other children living in the household and 22.7% of the maternal participants had not completed secondary school.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Attrition Status and Follow-up

Of the 2222 participating families, 799 families (36.0%) had complete follow-up, 137 (6.2%) were withdrawn, 308 (13.8%) were lost to follow-up and 978 families (44.0%) were partial or non-responders (Figure 1). There was no significant difference in the proportion of households with complete or incomplete follow-up across the 5 recruitment waves (p=0.20).

Insert Figure 1 about here.

Follow-up consisted of a total of 11,908.3 PYs from birth to seven years of age, with a mean 5.3 PYs per child (range 0-7). Automated linkage and manual searching by the state health department data linkage unit found records in QHAPDC for 97.1% of the child participants (n=2245), including their birth record.

Baseline Demographic, Household and Child Characteristics by Attrition Status

Baseline demographic and household characteristics differed significantly between participants who had complete follow up compared to those families who had incomplete follow up. Attrition families were more likely to have lower levels of maternal education, lower maternal age, have higher rates of maternal smoking, be single parent households, have more children residing in the home, have lower gross household incomes and live in rental or government boarding housing (all p<0.0001). However there were no differences in the proportion of child gender for active families compared to attrition families (Table 1).

Insert Table 1 about here.

These baseline characteristic differences were similar for two of the three attrition groups, lost to follow-up and partial/non-responders. However, those families in the withdrawn group were more similar in baseline characteristics to the families in the complete follow up group, with the exception of maternal age, marital status and gross household income (Table 1). Of note, families in the withdrawn group were more likely to have male children than those families who fully participated (p<0.01)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Prevalence Rates of Childhood Injury by Attrition Status

The total cohort had an overall child injury rate of 2.59/10PYs, similar to the child injury rate for those families with complete follow up (2.60/10PYs) and those families who were lost to follow-up or partial/non-responders (Table 2). However, families in the withdrawn group had significantly higher rates of child injury compared to families with complete follow up with an unadjusted RR=1.32 (95% CI 1.03-1.68). While there was little reduction in the effect size after adjusting for co-variates, the difference in child injury rates for families who withdrew compared to families with complete follow up was no longer statistically significant, with an adjusted RR=1.24 (95% CI 0.96-1.58). (Table 2).

Insert Table 2 about here.

Relationship between Baseline Sociodemographic Variables (exposure) and Child Injury Episodes (outcome) by Attrition Status

For the total cohort, rates of child injury increased as the level of maternal education, maternal age and gross household income decreased (Table 3). Rates of child injury were higher in single parent households, families living in government or boarding house accommodation and for families with male children. This distribution was similarly reflected in both the families with complete follow up and those lost to attrition, with no statistically significant differences in the demographically stratified rates of child injury when comparing overall attrition status ($p>0.05$) (Table 3).

Insert Table 3 about here.

While the direction of the demographically stratified child injury rates was similar for all attrition groups, the magnitude of this difference was greater in the withdrawn group for a number of baseline factors. Child injury rates for families that withdrew were higher than families with complete follow up for all baseline factors, but most pronounced with respect to child gender, maternal education, maternal smoking, number of children living in the household and home ownership ($p<0.05$).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

DISCUSSION

This study examined the impact of attrition in child injury research on the prevalence of child injury outcomes and the association between sociodemographic variables and injury (exposure-outcome relationship). There were three key findings. First, participants with incomplete follow up, as a group, differed from complete responders across most sociodemographic characteristics and there were also differences in these characteristics across attrition types (within the incomplete follow up group). Second, despite these systematic differences, there were no statistically significant impacts on the overall injury prevalence estimates and risk factor effect estimates, with the exception of the withdrawn group. The withdrawn group differed least, however, from the complete group in sociodemographic characteristics. Third, despite these systemic differences, the direction and nature of the sociodemographic-child injury relationship did not vary across the attrition groups.

The finding that there were systemic differences in the baseline social, maternal, and household characteristics of those with complete follow up versus those lost to attrition is consistent with previous research in this area, with families categorised as having lower socioeconomic status less likely to remain as active participants in the study.[3 5 9 10] Systematic differences in attrition were also evident across attrition groups, as has been demonstrated in previous research.[9 10] Interestingly, whilst families from the 'lost to follow up' and 'partial/non responder' groups were more likely to be of lower socioeconomic status, our withdrawn participants were more similar in their baseline characteristics to participants with complete follow up. It may be that families of higher sociodemographic status felt more confident to actively withdraw from this research or that there are unmeasured characteristics in these families that lead to higher rates of withdrawal. Further research is warranted given the relatively small sample size of this withdrawn group.

Consistent with previous research,[3 4 27] it was anticipated that systemic differences in attrition would result in differences in the prevalence of child injury outcomes across our groups (attrition-child injury), particularly as the relationship between lower sociodemographic factors and child injury has been well documented in the literature.[28-31] However, there were no statistically significant impacts of

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

this differential attrition on the overall injury prevalence estimates and risk factor effect estimates. The only apparent suggestion of difference (albeit with the observed stable effect losing significance with inclusion of multiple variables in the model) was in the withdrawn group, despite this group differing least from the "complete follow up" group for all measures of SES, compared to the other attrition categories. It should be noted though that the withdrawn group was small in absolute terms (i.e. <5% of the total cohort) and had the group been larger, it could have impacted on the overall prevalence of the outcome. This suggests that reasons for attrition may need to be examined in epidemiological research, as a high proportion of withdrawals may indicate the potential for bias.

Importantly, the systemic differences in attrition also had little impact on the direction of the relationship between sociodemographic factors and child injury, with lower socioeconomic status families having higher rates of injury across all the attrition groups. This finding adds to a growing body of evidence indicating that selective attrition does not necessarily impact the relationship between a range of exposure-outcome measures.[3 4 10] With follow up rates for cohort studies regularly being reported in the 30-70% range,[32] it is promising that these results suggest exposure-outcome findings from studies with large amounts of attrition may not need to be interpreted with the degree of caution currently applied.

A key strength of this study was the mixed active-passive nature of the cohort design. Prior to the first published results from this study, there had been few birth cohort studies that employed data linkage methodology in combination with active participant contact (survey follow up) to ascertain comprehensive exposure and outcome data. Thus, this is one of the few studies capable of examining with rigour the questions we have addressed in this research. The key limitation is the relatively small sample size. This was particularly evident in relation to our main finding that it was in the group of participants who withdrew from the study that we found a bias. The adjusted model removed significance of the main finding even though there was little change in the point estimate (from 1.32 to 1.24) and could represent a type 2 error. The second limitation relates to the nature of the study sample, the determinant variables measured, the injury outcome focus of the analysis and the follow up protocols used. While the research results could be expected to apply to similar circumstances, it is

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

unknown how validly the study findings should be applied beyond the study group we examined. It is possible that in studies with a demanding follow up protocol, participant attrition may be higher in people whose determinant variables are already affecting outcomes, and hence in these studies attrition may be more likely to be associated with a biased effect estimate.

The findings of this research provide support for relaxing one of the most challenging expectations of epidemiological studies; that is, complete participant follow up. Even with considerable attrition, when proportions of participants who withdraw from the study are low, attrition bias is unlikely to affect either the population prevalence estimate or measures of association with key study variables.

ACKNOWLEDGMENTS

The research reported in this publication is part of the Environments for Healthy Living (EFHL): Griffith Study of Population Health (Australian and New Zealand Clinical Trials Registry: ACTRN12610000931077). The EFHL project was conceived by Professor Rod McClure, Dr Cate Cameron, Professor Judy Searle and Professor Ronan Lyons. We are thankful for the contributions of the Project Manager, Rani Scott, and the current and past Database Managers. We gratefully acknowledge the administrative staff, research staff, and the hospital antenatal and birth suite midwives of the participating hospitals for their valuable contributions to the study, in addition to the expert advice provided by Research Investigators throughout the project.

FUNDING

This study was funded by an Australian Research Council Discovery Grant [DP110105423]. Core funding to support EFHL is provided by Griffith University. Dr Cameron was supported by a Public Health Fellowship (ID 428254) from the NHMRC.

COMPETING INTERESTS

The authors do not have any competing interests in relation to this research.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

REFERENCES

1. Western M, Haynes M, Baffour B, et al. The Case for a New Australian Birth Cohort Study: Institute for Social Science Research, University of Queensland 2015.
2. Ahern K, Le Brocque R. Methodological Issues in the Effects of Attrition: Simple Solutions for Social Scientists. *Field Methods* 2005;17(1):53-69.
3. Wolke D, Waylen A, Samara M, et al. Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British journal of psychiatry : the journal of mental science* 2009;195(3):249-56. doi: 10.1192/bjp.bp.108.053751 [published Online First: 2009/09/02]
4. Osler M, Kriegbaum M, Christensen U, et al. Loss to follow up did not bias associations between early life factors and adult depression. *J Clin Epidemiol* 2008;61(9):958-63. doi: 10.1016/j.jclinepi.2007.11.005 [published Online First: 2008/05/23]
5. Greene N, Greenland S, Olsen J, et al. Estimating bias from loss to follow-up in the Danish National Birth Cohort. *Epidemiology* 2011;22(6):815-22. doi: 10.1097/EDE.0b013e31822939fd [published Online First: 2011/09/16]
6. Ng SK, Scott R, Scuffham PA. Contactable Non-responders Show Different Characteristics Compared to Lost to Follow-Up Participants: Insights from an Australian Longitudinal Birth Cohort Study. *Matern Child Health J* 2016;20(7):1472-84.
7. Howe LD, Tilling K, Galobardes B, et al. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. *Epidemiology* 2013;24(1):1-9. doi: 10.1097/EDE.0b013e31827623b1 [published Online First: 2012/12/06]
8. Aylward GP, Hatcher RP, Stripp B, et al. Who goes and who stays: subject loss in a multicenter, longitudinal follow-up study. *J Dev Behav Pediatr* 1985;6(1):3-8.
9. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? *Aust N Z J Public Health* 2006;30(4):353-61. [published Online First: 2006/09/08]
10. Powers J, Tavener M, Graves A, et al. Loss to follow-up was used to estimate bias in a longitudinal study: a new approach. *J Clin Epidemiol* 2015;68(8):870-6. doi: 10.1016/j.jclinepi.2015.01.010 [published Online First: 2015/02/24]
11. Cornish R, Tilling K, Boyd A, et al. Using linkage to electronic primary care records to evaluate recruitment and nonresponse bias in the Avon Longitudinal Study of Parents and Children. *Epidemiology* 2015;26(4):e41-2. doi: 10.1097/EDE.0000000000000288 [published Online First: 2015/04/04]
12. Jourdan C, Bayen E, Bahrami S, et al. Loss to follow-up and social background in an inception cohort of patients with severe traumatic brain injury: results from the PariS-TBI study. *The Journal of head trauma rehabilitation* 2016;31(3):E42-E48.
13. Tin Tin S, Woodward A, Ameratunga S. Estimating bias from loss to follow-up in a prospective cohort study of bicycle crash injuries. *Inj Prev* 2014;20(5):322-9. doi: 10.1136/injuryprev-2013-040997
14. Holavanahalli RK, Lezotte DC, Hayes MP, et al. Profile of patients lost to follow-up in the Burn Injury Rehabilitation Model Systems' longitudinal database. *J Burn Care Res* 2006;27(5):703-12. doi: 10.1097/01.BCR.0000238085.87863.81
15. Blaha RZ, Arnett AB, Kirkwood MW, et al. Factors influencing attrition in a multisite, randomized, clinical trial following traumatic brain injury in adolescence. *J Head Trauma Rehabil* 2015;30(3):E33-40. doi: 10.1097/HTR.0000000000000059
16. Corrigan JD, Harrison-Felix C, Bogner J, et al. Systematic bias in traumatic brain injury outcome studies because of loss to follow-up. *Arch Phys Med Rehabil* 2003;84(2):153-60. doi: 10.1053/apmr.2003.50093
17. Gabbe BJ, Cleland H, Watterson DM, et al. Long term outcomes data for the Burns Registry of Australia and New Zealand: Is it feasible? *Burns : journal of the International Society for Burn Injuries* 2015;41(8):1732-40. doi: 10.1016/j.burns.2015.09.005

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

18. Wade SL, Gerry Taylor H, Yeates KO, et al. Long-term parental and family adaptation following pediatric brain injury. *Journal of pediatric psychology* 2006;31(10):1072-83. doi: 10.1093/jpepsy/jsj077
19. Krellman JW, Kolakowsky-Hayner SA, Spielman L, et al. Predictors of follow-up completeness in longitudinal research on traumatic brain injury: findings from the National Institute on Disability and Rehabilitation Research traumatic brain injury model systems program. *Arch Phys Med Rehabil* 2014;95(4):633-41. doi: 10.1016/j.apmr.2013.10.016
20. de Groot S, Haisma JA, Post MW, et al. Investigation of bias due to loss of participants in a Dutch multicentre prospective spinal cord injury cohort study. *J Rehabil Med* 2009;41(5):382-9. doi: 10.2340/16501977-0346
21. Neuner B, Fleming M, Born R, et al. Predictors of loss to follow-up in young patients with minor trauma after screening and written intervention for alcohol in an urban emergency department. *Journal of studies on alcohol and drugs* 2007;68(1):133-40.
22. Lin C-YA, DeMatteo CA. The Challenges of Loss to Follow-up in Longitudinal Pediatric Acquired Brain Injury Research: One Research Team's Experiences. 2013
23. Calderwood L, Lessof C. Enhancing longitudinal surveys by linking to administrative data. *Methodology of longitudinal surveys* 2009:55-72.
24. Cameron CM, Scuffham PA, Spinks A, et al. Environments for Healthy Living (EFHL) Griffith birth cohort study: background and methods. *Matern Child Health J* 2012;16(9):1896-905. doi: 10.1007/s10995-011-0940-4 [published Online First: 2012/02/09]
25. Cameron CM, Scuffham PA, Shibl R, et al. Environments For Healthy Living (EFHL) Griffith birth cohort study: characteristics of sample and profile of antenatal exposures. *BMC Public Health* 2012;12:1080. doi: 10.1186/1471-2458-12-1080 [published Online First: 2012/12/18]
26. State of Queensland (Queensland Health). Queensland Data Linkage Framework. Brisbane, 2016.
27. Mars B, Cornish R, Heron J, et al. Using Data Linkage to Investigate Inconsistent Reporting of Self-Harm and Questionnaire Non-Response. *Archives of suicide research : official journal of the International Academy for Suicide Research* 2016;20(2):113-41. doi: 10.1080/13811118.2015.1033121 [published Online First: 2016/01/21]
28. Hong J, Lee B, Ha EH, et al. Parental socioeconomic status and unintentional injury deaths in early childhood: consideration of injury mechanisms, age at death, and gender. *Accident; analysis and prevention* 2010;42(1):313-9. doi: 10.1016/j.aap.2009.08.010 [published Online First: 2009/11/06]
29. Kim M-H, Subramanian SV, Kawachi I, et al. Association between childhood fatal injuries and socioeconomic position at individual and area levels: a multilevel study. *J Epidemiol Community Health* 2007;61:135-40. doi: doi:10.1136/jech.2006.047738
30. Laursen B, Nielsen JW. Influence of sociodemographic factors on the risk of unintentional childhood home injuries. *European journal of public health* 2008;18(4):366-70. doi: 10.1093/eurpub/ckn034 [published Online First: 2008/06/03]
31. Myhre MC, Thoresen S, Grogaard JB, et al. Familial factors and child characteristics as predictors of injuries in toddlers: a prospective cohort study. *BMJ open* 2012;2(2):e000740. doi: 10.1136/bmjopen-2011-000740 [published Online First: 2012/03/10]
32. Gustavson K, von Soest T, Karevold E, et al. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC public health* 2012;12:918. doi: 10.1186/1471-2458-12-918

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Figure 1 Flow diagram of EFHL cohort attrition status, Queensland participants 2006-2010 (N=2222 Households and N=2245 Children).

[attached file]

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1 Distribution of baseline maternal, household and child characteristics, stratified by follow-up status and types of attrition.

	Follow-up			Attrition by type [a]		
	Complete	Incomplete		Withdrawn	Lost to Follow-up	Partial/Non Responders
	N (%)	N (%)	p-value	N (%)	N (%)	N (%)
MATERNAL CHARACTERISTICS	**799**	**1423**		**137**	**308**	**978**
Maternal level of education *Missing=12*						
Tertiary degree	192 (24.1)	178 (12.6)	***	24 (17.5)	34 (11.0) ***	120 (12.4) ***
Secondary school or trade	490 (61.6)	849 (60.0)		89 (65.0)	182 (59.3)	578 (59.6)
Not complete school	114 (14.3)	387 (27.4)		24 (17.5)	91 (29.6)	272 (28.0)
Maternal age *Missing=3*						
35+ years	201 (25.2)	228 (16.1)	***	17 (12.4) ***	53 (17.2) ***	158 (16.2) ***
30-34 years	253 (31.7)	308 (21.7)		37 (27.0)	56 (18.2)	215 (22.0)
25-29 years	228 (28.6)	416 (29.3)		48 (35.0)	83 (26.9)	285 (29.2)
<25 years	116 (14.5)	469 (33.0)		35 (25.6)	116 (37.7)	318 (32.6)
Smoking during pregnancy *Missing=9*						
No	640 (80.1)	949 (67.1)	***	102 (74.5)	187 (60.9) ***	660 (68.0) ***
Yes	159 (19.9)	465 (32.9)		35 (25.5)	120 (39.1)	310 (32.0)
HOUSEHOLD CHARACTERISTICS	**799**	**1423**		**137**	**308**	**978**
Marital status *Missing=18*						
Two parent family	732 (92.1)	1161 (82.4)	***	106 (77.4) ***	247 (81.2) ***	808 (83.5) ***
Sole parent family	63 (7.9)	248 (17.6)		31 (22.6)	57 (18.8)	160 (16.5)
Other children in household *Missing=14*						
None	299 (37.6)	471 (33.4)	***	53 (39.0)	97 (31.8) ***	321 (33.0) ***
1-2 children	431 (54.2)	731 (51.7)		70 (51.5)	154 (50.5)	507 (52.2)
3+ children	65 (8.2)	211 (14.9)		13 (9.5)	54 (17.7)	144 (14.8)
Household income *Missing=354*						
Highest quintile	170 (24.0)	175 (15.1)	***	17 (15.2) *	30 (12.1) ***	128 (16.0)***
4th quintile	154 (21.7)	206 (17.8)		31 (27.7)	31 (12.4)	144 (18.1)
3rd quintile	164 (23.1)	211 (18.2)		21 (18.7)	42 (16.9)	148 (18.6)
2nd quintile	129 (18.2)	259 (22.3)		19 (17.0)	65 (26.1)	175 (21.9)
Lowest quintile	92 (13.0)	308 (26.6)		24 (21.4)	81 (32.5)	203 (25.4)
Home ownership *Missing=20*						
Own/mortgage	427 (53.6)	437 (31.1)	***	62 (45.3)	49 (16.2) ***	326 (33.7) ***
Rent	297 (37.2)	740 (52.7)		57 (41.6)	203 (67.2)	480 (49.7)
Government/Boarding house	73 (9.2)	228 (16.2)		18 (13.1)	50 (16.6)	160 (16.6)
CHILD CHARACTERISTICS	**808**	**1437**		**137**	**309**	**991**
Gender *Missing = 49*						
Female	400 (50.1)	670 (48.0)	ns	51 (37.5) **	151 (49.7)	468 (48.9)
Male	399 (49.9)	727 (52.0)		85 (62.5)	153 (50.3)	489 (51.1)

Maternal and Household N=2222; Children N=2245 (includes twins)

[a] Each Drop-out participant group compared to the complete participant group as the reference, using chi-squared test.

*p<0.05 ** p<0.01 ***p <0.0001 ns= not significant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2 Number of injuries, injury rates per 10 person-years (PYs) and Injury Rate Ratios (RRs) with 95% Confidence Intervals (CIs) for the child participants, stratified by follow-up status and types of attrition.

Child Participants	N (%)	Sum Injuries	Injury Rates /10 PYs	Unadjusted RR (95% CI) [a]	Adjusted RR (95% CI) [a,b]
Total cohort	2245 (100.0)	941	2.59	0.99 (0.88-1.13)	0.97 (0.85-1.10)
Complete follow-up	808 (36.0)	340	2.60	Reference	Reference
Incomplete follow-up	1437 (64.0)	601	2.58	0.99 (0.87-1.13)	0.95 (0.82-1.09)
Withdrawn	137 (6.1)	80	3.48	1.32 (1.03-1.68) *	1.24 (0.96-1.58)
Lost to follow-up	309 (13.8)	128	2.38	0.89 (0.72-1.09)	0.81 (0.65-1.01)
Partial/Non-responders	991 (44.1)	393	2.51	0.98 (0.85-1.13)	0.94 (0.81-1.10)

[a] Poisson regression, offset = log of length of follow-up time of child.

[b] Adjusted for maternal age, maternal education, maternal smoking status, number of children in household, family marital status, household income, housing status and child gender in two step model. Final model included maternal age, maternal education level and child gender

*p<0.05

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3 Child injury rates per 10 person-years (PYs) for the total cohort and stratified by follow-up status and types of attrition for baseline maternal, household and child characteristics.

	Total Cohort	Follow-up			Attrition by type [a]		
		Complete	Incomplete		Withdrawn	Lost to Follow-up	Partial/Non Responders
	/10 PYs	/10 PYs	/10 PYs	p-value	/10 PYs	/10 PYs	/10 PYs
MATERNAL CHARACTERISTICS							
Maternal level of education *Missing=12*							
Tertiary degree	2.03	2.12	1.92	ns	3.06 *	1.72	1.74
Secondary school or trade	2.70	2.84	2.62		3.66	2.28	2.57
Not complete school	2.73	2.43	2.83		3.23	2.87	2.77
Maternal age *Missing=3*							
35+ years	2.02	1.83	2.18	ns	2.83	2.76	1.91
30-34 years	2.67	2.86	2.52		3.93	1.50	2.55
25-29 years	2.27	2.35	2.23		2.93	1.69	2.27
<25 years	3.27	3.76	3.14		4.12	3.13	3.04
Smoking during pregnancy *Missing=9*							
No	2.49	2.48	2.50	ns	3.25 *	2.55	2.36
Yes	2.84	3.10	2.75		4.17	2.15	2.84
HOUSEHOLD CHARACTERISTICS							
Marital status *Missing=18*							
Two parent family	2.45	2.53	2.41	ns	3.36	2.33	2.30
Sole parent family	3.44	3.47	3.43		3.90	2.66	3.64
Other children in household *Missing=14*							
None	2.70	2.84	2.61	ns	3.09 *	2.66	2.51
1-2 children	2.52	2.51	2.53		3.50	2.30	2.46
3+ children	2.52	1.99	2.68		4.21	2.17	2.73
Household income [b] *Missing=354*							
Highest quintile	2.03	2.42	1.67	ns	2.04	1.52	1.65
4th quintile	2.55	2.93	2.27		2.79	1.95	2.24
3rd quintile	2.45	2.33	2.55		3.57	2.92	2.27
2nd quintile	2.54	2.32	2.65		3.74	2.32	2.66
Lowest quintile	2.92	3.53	2.74		3.15	2.36	2.85
Home ownership *Missing=20*							
Own/mortgage	2.41	2.53	2.29	ns	3.68 *	1.47	2.16
Rent	2.46	2.48	2.46		2.54	2.50	2.43
Government/Boarding house	3.55	3.39	3.60		5.85	2.93	3.57
CHILD CHARACTERISTICS							
Gender *Missing = 49*							
Female	2.18	2.21	2.16	ns	3.13 *	1.89	2.15
Male	3.00	2.87	3.07		3.74	2.93	2.99

Children N=2245 (includes twins)

[a] Each Drop-out participant group compared to the complete participant group as the reference, using univariate Poisson regression with time offset.

[b] Gross annual household incomes standardised in AUD$2010 values

*p<0.05 ** p<0.01 ***p <0.0001 ns=not significant

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

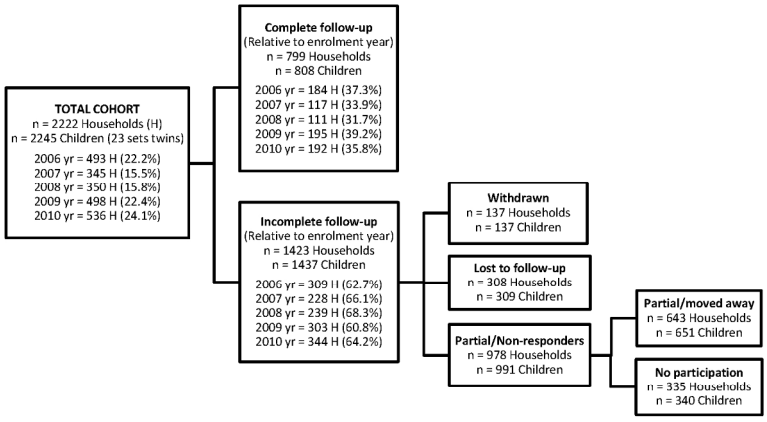

Figure 1 Flow diagram of EFHL cohort attrition status, Queensland participants 2006-2010 (N=2222 Households and N=2245 Children).

297x210mm (200 x 200 DPI)

For peer review only

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	PAGE
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-6
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	6
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	Na
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-8
Bias	9	Describe any efforts to address potential sources of bias	**Paper about bias**
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9
		(*b*) Describe any methods used to examine subgroups and interactions	9
		(*c*) Explain how missing data were addressed	9
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	**Paper about bias from loss to follow-up**
		(*e*) Describe any sensitivity analyses	

Continued on next page

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Figure 1
		(b) Give reasons for non-participation at each stage	Figure 1
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	Table 1
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	9-10
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	Table 3
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2, page 11
		(*b*) Report category boundaries when continuous variables were categorized	Table 1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11
Discussion			
Key results	18	Summarise key results with reference to study objectives	12-13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13-14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-14
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml