Longitudinal and Life Course Studies • vol xx • no xx • 1–14
 • Paper received /Revised /Published online October 2020
Online ISSN 1757-9597 • https://doi.org/10.1332/175795920X16000977636901

RESEARCH ARTICLE

Characterising attrition from childhood to adulthood in a 20-year cohort: which baseline factors are influential, and can bias be corrected?

Petr Otahal, potahal@utas.edu.au
Leigh Blizzard, Leigh.Blizzard@utas.edu.au
David W. Hosmer, hosmer@schoolph.umass.edu
University of Tasmania, Australia
and University of Vermont, USA

Jim Stankovich, jim.stankovich@monash.edu
Monash University, Australia
and University of Tasmania, Australia

Alison Venn, alison.venn@utas.edu.au
University of Tasmania, Australia

Attrition is common in longitudinal studies and can lead to bias when the missingness pattern affects the distributions of analysed variables. Characterisation of factors predictive of attrition is vital to longitudinal research. Few studies have investigated the factors predictive of attrition from childhood cohorts with large-scale loss to follow-up. Methods to remove potential bias are available and have been well studied in scenarios of short intervening periods between contact and follow-up. Less is known about the performance of such techniques when there is a large initial loss of participants after a long intervening period. The Australian Schools Health and Fitness Survey (ASHFS) was conducted in 1985 when participants were school children aged 7–15 years. The first follow-up occurred 20 years later with substantial loss of participants: 80% were traced, 61% enrolled and provided brief questionnaire information, 47% provided more extensive questionnaire information and 28% attended clinics. Factors associated with attrition were examined and two common techniques, multiple imputation (MI) and inverse probability weighting (IPW) were used to determine the potential for correcting the bias in the estimate of the association between self-rated fitness and BMI in childhood. Attrition from childhood to adulthood was found to be influenced by the same factors that operate in adult cohorts: lower education, lower socio-economic position and male sex. Attrition patterns varied by the stage of follow-up. Estimated childhood associations biased by adulthood attrition were able to be corrected using MI, but IPW was unsuccessful due to a lack of completely observed informative variables.

Key words follow-up • attrition • bias • imputation • inverse probability weighting • validity

Key messages

- Education, SEP and male sex are key predictors of attrition in a cohort from childhood to adulthood.
- Inverse probability weighting for missing data performs poorly without key complete variables.
- Multiple imputation was able to correct biased estimates even when attrition was greater than 70%.

To cite this article: Otahal, P., Blizzard, L., Hosmer, D., Stankovich, J. and Venn, A. (2020) Characterising attrition from childhood to adulthood in a 20-year cohort: which baseline factors are influential, and can bias be corrected?, *Longitudinal and Life Course Studies*, vol xx, no xx, 1–14, DOI: 10.1332/175795920X16000977636901

Introduction

Longitudinal studies are valuable in providing epidemiological evidence to support inferences about causal relationships, but most suffer from participant attrition. Attrition is usually a gradual loss of study participants during follow-up and is typically a non-random process: specific patterns arise due to selective non-participation of individuals. Non-participation is thus a potential source of selection bias for estimates of associations under study. The bias occurs when the associations of interest differ for missing participants and those remaining in the study (Greenland, 1977). Valid techniques exist to remove bias if the missing data are missing at random (MAR), or even if they are missing not at random (MNAR) (Rubin, 1976), but they rely on untestable assumptions about the nature of the missing data mechanism. Nevertheless, identifying factors informative of dropout is a prerequisite initial step for any analysis of studies with missing data and an important procedural check to ensure robustness of inference in studies that suffer from attrition.

Attrition and non-response in surveys has been extensively studied, particularly in household panel surveys and birth cohort studies (Fitzgerald et al, 1998; Gustavson et al, 2012; Brick, 2013; Mostafa and Wiggins, 2015). The patterns of attrition vary from study to study, but poor education, lower socio-economic background and male sex have been shown to be associated with non-participation at follow-up. A number of studies examining attrition have been conducted for participants recruited in childhood or adolescence (Siddiqui et al, 1996; Burkam and Lee, 1998; Eerola et al, 2005; Olsen, 2005; Cumming and Goldstein, 2016) and, in those that have, survey participants were recontacted within short (one to three year) follow-up periods. This frequency of contact results in lower levels of attrition through better cohort maintenance and may alter the mix of factors associated with attrition. There are few studies examining attrition with a long delay between baseline and follow-up (Badawi et al, 1999), and none in which participants were enrolled during childhood with follow-up conducted in adulthood. Consequently, there is little understanding of which factors are associated with attrition in such studies and whether techniques for handling missing data in other scenarios are applicable to this setting.

We had the opportunity to investigate this issue using an Australian cohort enrolled in childhood in 1985 when participants were school children aged 7–15 years (Pyke, 1987). This cohort was initially recruited for a cross-sectional survey with no intention of follow-up and no infrastructure to maintain contact with the cohort members. Follow-up was initiated approximately 17 years after the original enrolment with the first follow-up measurements occurring approximately 20 years post baseline. The cohort had a high rate of non-participation, particularly in intensive clinical examinations, during the first follow-up. The purpose of this study is to describe attrition from the original cohort at each of four stages of the first follow-up – contact tracing, study enrolment, response to questionnaires and attendance at clinics for physical examination – and investigate the factors associated with that attrition. Additionally, we demonstrate the bias that arises in a baseline association when it is estimated using the data available for participants at each of the four stages of follow-up. We compare the performance of two common missing data analysis techniques (MI and IPW) in removing this bias at final stage of follow-up, the stage with the highest participant attrition.

Methods

Subjects

Baseline data collection commenced with the Australian Schools Health and Fitness Survey (ASHFS) of 7–15 year-old schoolchildren conducted in 1985 (Pyke, 1987). The sample for the ASHFS survey was selected by two-stage probability sampling: participating schools from every state and territory of Australia were selected at the first stage with probability proportional to enrolment size, followed by age- and sex-stratified random sampling of students within schools at the second stage. The initial ASHFS cohort comprised 8,498 children from 109 schools. All participants were involved in a battery of measurements of anthropometry and field tests of physical fitness. Participants aged 9, 12 or 15 years completed a series of technical tests of health and fitness, and a subset then completed laboratory tests. A questionnaire was administered in small supervised groups to all 6,559 children aged nine years or over. It collected information on demographic and family characteristics, history of involvement in exercise and sport, health-related behaviours, and attitudes to health-related factors.

The first Childhood Determinants of Adult Health (CDAH) follow-up of the original ASHFS cohort occurred during 2004–06, approximately 20 years after the initial ASHFS study. CDAH has now had several follow-up phases but this study is limited to the first follow-up. Contact with the initial survey respondents was re-established during 2002–03 after tracing as many of the original members as possible using the only identifying information available (participants' name, date of birth, sex and school attended). Follow-up questionnaires were administered, and clinics were conducted during 2004–06. There were four distinct stages of data collection as part of this first follow-up:

1 Tracing: initial tracing of ASHFS participants involving record searching and family and peer contact.
2 Enrolment: agreement and consent to participate in the CDAH follow-up was obtained, and a short self-report questionnaire was completed.
3 Questionnaire: information was obtained by completion of a mailed questionnaire in the first instance. Non-responders were asked to complete and return a

shortened version of the questionnaire that was mailed to them. In the case of further non-response, they were asked to participate in a short Computer Assisted Telephone Interview (CATI).

4 Clinic: physical and psychometric tests were administered, and fasting blood samples were taken for chemical analysis.

The tracing process commenced with linkage of ASHFS participants on name and date of birth to the National Death Index and Australian Electoral Commission electoral rolls. Electoral rolls were an important data source because it is compulsory by law for all eligible Australian citizens to enroll and vote in federal, state, and local government elections, by-elections, and referendums. A number of periods were used: the current roll available at the time of tracing (2002–03), historical rolls from 1985 that were searched for persons of the same surname as the ASHFS participants living in the vicinity of the schools in that year, and historical rolls corresponding to the year in which the ASHFS participants turned 18 years of age. Where name and date of birth were able to be matched and the person was not deceased, the information from these rolls was used to deduce the most recent contact address for participants of the original ASHFS study or that of their parents. Where participants could not be located using Electoral Roll information, the electronic white pages telephone directory and the Australia on Disc (United Directory Systems, 2018) database were searched. Further tracing was performed by contacting friends on the now discontinued School Friends website that was active during 2000–05. This was a social networking platform connecting former classmates.

Measurements

The measurements of anthropometry and field tests of physical fitness in 1985 were conducted on the entire cohort. They included measurement of height, weight, waist circumference, 1.6 km timed run and 50 m timed sprint. Height was measured using a KaWe Height tape or rigid metric measuring tape and plastic set square, and weight was measured using calibrated beam or medical spring scales. Field tests were conducted after a thorough warm-up, and included a timed 1.6 km rum on a round or oval 200 m or 400 m track and a timed 50 m sprint across the wind in groups of four students. Additionally, age, sex, school type (independent/Catholic/public), and name (given name and surname) were recorded for the entire cohort. Body mass index (BMI) was derived from the measured height and weight variables as the weight in kilograms divided by the square of height in metres.

Questionnaire-derived variables for 9–15 year-old children included classification by socio-economic position (SEP), by rurality and by teacher-rated scholastic ability. SEP was derived from postcode using the Australian Bureau of Statistics (ABS) Socio-economic Index of Relative Disadvantage (SEIFA) classified in quarters of the distribution. Rurality was derived from postcode of residence using the ABS 1986 Census Australian Standard Geographical Classification (ABS catalogue 1216.0). The categories include major urban (population $>$ 100,000), other urban (1,000–99,999), bounded locality (200–999) and rural ($<$200), and with the latter two categories combined for analysis. Scholastic ability was a teacher-assessed rating given by the school for each student on a five-point Likert scale ranging from Poor to Excellent. A further questionnaire-derived variable: self-rated fitness at baseline was derived from responses

Table 1: Characteristics of participants at baseline (ASHFS 1985)

	Percentage or Mean (SD)	n/N or N
Age group		
7–9 years	34.2	2,909/8,498
10–12 years	34.7	2,946/8,498
13–15 years	31.1	2,643/8,498
Sex		
Female	49.3	4,191/8,498
Male	50.7	4,307/8,498
School type		
Government (public)	75.0	6,375/8,498
Catholic	19.7	1,673/8,498
Independent (private)	5.3	450/8,498
Socio-economic position		
1st quarter (high)	23.7	1,490/6,299
2nd quarter	28.6	1,800/6,299
3rd quarter	38.5	2,427/6,299
4th quarter (low)	9.2	582/6,299
Rurality		
Major urban	62.7	3,982/6,346
Other urban	18.2	1,158/6,346
Rural	19.0	1,206/6,346
Teacher-rated scholastic ability		
Excellent	9.3	743/7,961
Above average	27.8	2,211/7,961
Average	41.2	3,279/7,961
Below average	16.8	1,336/7,961
Poor	4.9	392/7,961
Self-rated fitness		
Not as fit as most	11.9	761/6,403
About average fitness	67.2	4,302/6,403
Fitter than most	20.9	1,340/6,403
BMI (kg/m^2)	18.23 (2.89)	8,491
Height (cm)	146 (5.43)	8,492
Weight (kg)	39.9 (3.04)	8,496
Long run time (minutes)	9.25 (1.94)	7,876
Short run time (seconds)	9.21 (1.08)	8,069

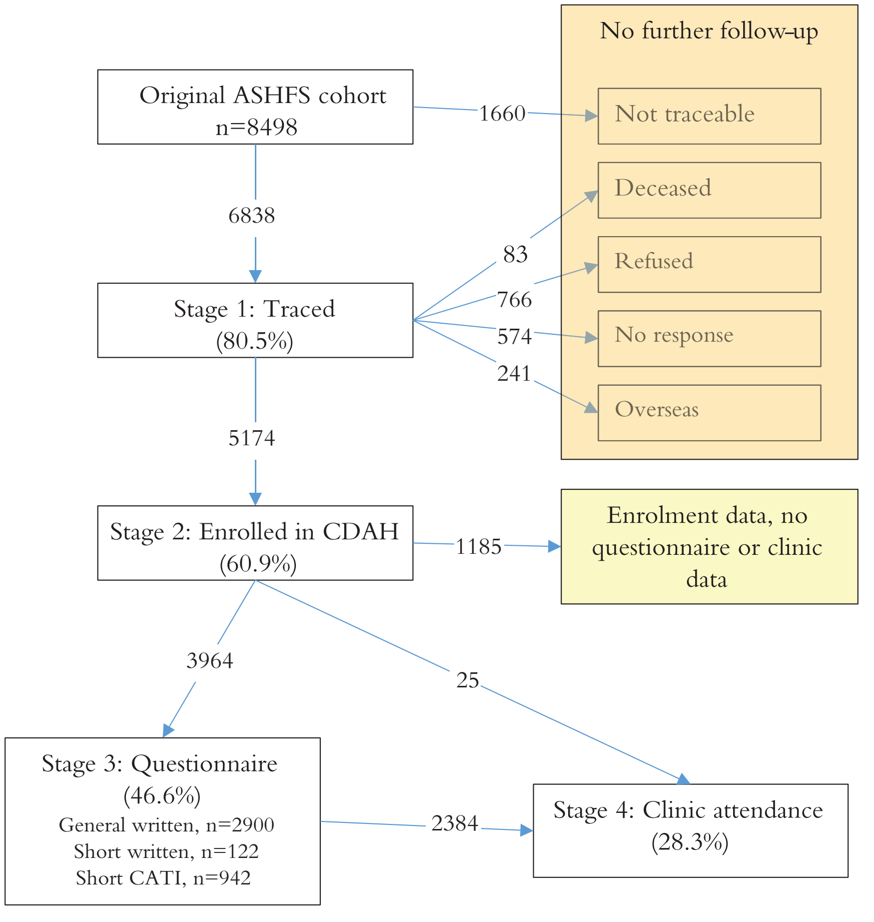
Notes: ASHFS: Australian Schools Health Fitness Survey; BMI: Body Mass Index; SD: standard deviation; n: number within group; N: total valid responses

to the question ‘How fit do you think you are compared to others of your age?’ with response categories ‘Fitter than most’, ‘About average fitness’, or ‘Not as fit as most’.

Analysis

Demographic data for the ASHFS cohort are presented as means and standard deviations for continuous variables, or percentages and frequencies for categorical

Figure 1: Cohort flow diagram of the baseline ASHFS cohort and four stages of the CDAH follow-up.



variables. Percentages of subjects remaining from the baseline ASHFS cohort at each of the four CDAH data collection stages are shown.

The association between BMI and self-rated fitness, both assessed at ASHFS, were estimated by linear regression. The analysis within this part of the study is restricted to the 6,398 children aged 9–15 years who completed the ASHFS questionnaire and additionally had complete data for self-rated fitness and BMI. The associations were re-estimated for those who participated at the four data collection stages of the initial CDAH follow-up. BMI in childhood varied with age and sex (Cole et al, 2007), and all models included linear covariates for these factors to adjust for their influence.

Two commonly used missing data techniques, inverse probability weighting (IPW) and multiple imputation (MI), were investigated in the linear regression of levels of baseline self-rated fitness regressed on baseline BMI. The participant data were: (1) those available for all participants at baseline, and (2) those available for participants at the Clinic stage of follow-up. Inverse probability weights were calculated from variables fully observed for 9–15-year-old participants (age, sex and school type) (Seaman and White, 2013). Following van Buuren (van Buuren and Groothuis-

Table 2: Response patterns for ASHFS participants at each of four stages of the first CDAH study follow-up

Variable	Level	Tracing*	Enrolment*	Questionnaire*	Clinic*
Age	07–09	78.5	58.6	43.1	26.0
	10–12	82.0	60.9	46.9	29.1
	13–15	81.0	63.2	50.3	30.2
Sex	Male	79.6	56.6	42.1	26.7
	Female	81.4	65.2	51.3	30.1
School type	Government	78.9	59.3	45.1	27.2
	Catholic	84.5	64.4	50.1	31.0
	Independent	87.6	68.9	55.8	34.9
Socio-economic position	1st quarter (high)	83.8	64.6	50.4	35.4
	2nd quarter	80.2	59.7	46.4	28.9
	3rd quarter	81.8	64.0	49.8	27.8
	4th quarter (low)	72.9	52.7	38.7	24.9
Rurality	Major urban	79.1	59.3	45.3	29.8
	Other urban	81.4	63.7	50.0	28.4
	Rural	87.3	69.2	54.8	30.1
Teacher-rated scholastic ability	Excellent	83.0	67.4	55.0	38.0
	Above average	84.3	67.1	53.1	33.9
	Average	81.1	61.1	47.3	27.9
	Below average	75.1	52.6	36.2	19.1
	Poor	71.4	44.4	25.0	12.5

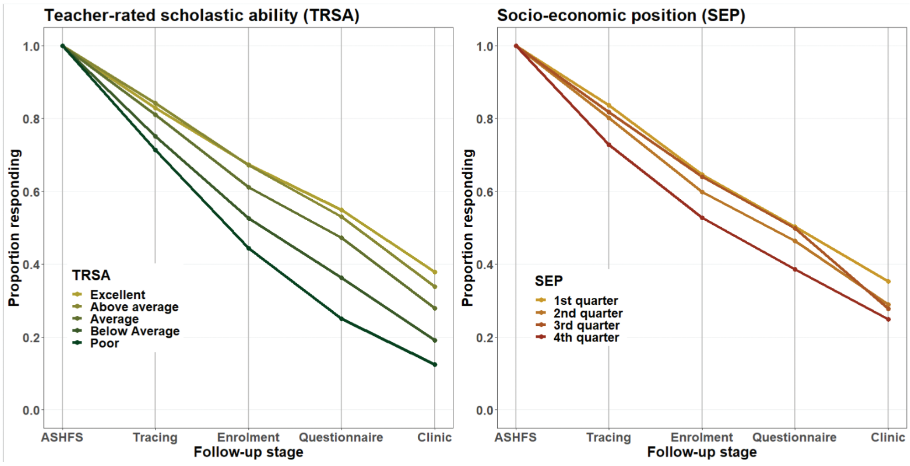
Notes: * Percentage of original ASHFS cohort shown within each level.
ASHFS: Australian Schools Health Fitness Survey; CDAH: Childhood Determinants of Adult Health.

Oudshoorn, 2011), covariates included in the MI models were those for variables in the analysis model (BMI, self-rated fitness, age and sex), variables associated with missingness (teacher-assessed scholastic ability, school type, rurality, SEP), and variables associated with BMI or self-rated fitness with correlation greater than 0.2 (height, weight, waist circumference, time for 1.6 km run, time for 50 m run). BMI was passively imputed from weight ÷ height^2 and was not included as a predictor of other variables. Fully Conditional Specification (chained equations) models were used for each analysis with 50 imputations following a burn-in of 10 iterations. Estimates from each imputed dataset were combined using Rubin's Rules (van Buuren, 2012). All analyses were performed in R (R Core Team, 2019), with inverse probability weights applied using the *survey* package and multiple imputation conducted using the *mice* package (van Buuren and Groothuis-Oudshoorn, 2011).

Results

The original 1985 ASHFS cohort comprised 8,498 children. Of those, 80.5% (6,838/8,498) were traced in 2002–03 and 60.9% (5,174/8,498) were enrolled in the CDAH study. Following enrolment, 46.6% (3,964/8,498) participated in survey questionnaires and 28.3% (2,409/8,498) attended clinics. Figure 1 shows participant numbers at each of these follow-up stages. For the most part, those lost through

Figure 2: Proportion of responders by category for teacher-rated scholastic ability and socio-economic position over follow-up stages.



attrition followed a monotonic pattern of missingness through the course of these four stages.

The characteristics of the ASHFS sample are shown in Table 1. Reflecting the socio-demographic characteristics of the nation in 1985, the sample comprised predominantly government (public) school participants from major urban areas. The level of missing data at baseline due to item non-response and study design (exclusion of 7/8-year-olds from questionnaires) varied from zero to 25.9%.

Table 2 shows the patterns of attrition from the original ASHFS cohort. Teacher-rated scholastic (TRSA) ability was the strongest predictor of attrition with an inverse association and a marked gradation of attrition from the highest to lowest levels. The disparity in attrition began at the Tracing stage and worsened through the stages of follow-up. Those rated as excellent had the least attrition (38.0% responding at Clinic) while those rated as poor had the highest attrition (12.5% responding at Clinic). SEP had an inverse association with attrition similar to that of TRSA: participants from high SEP areas (first quarter) had the least attrition (35.4% responding at Clinic) while those from low SEP areas (fourth quarter) had highest attrition (24.9% responding at Clinic). From the Clinic stage, attrition decreased monotonically with higher SEP. Male participants had high levels of attrition compared with female participants at the Enrolment and Questionnaire stages, but this difference was less pronounced at other stages. The type of school attended also characterised the level of attrition. Children from independent (private) schools had the lowest rates of attrition, those from Catholic schools had higher rates, and finally those from government (public) schools had the highest rates of attrition. Rurality was associated with attrition for the first three stages of follow-up with greater attrition for those participants living in more urban areas. This trend in rurality was not observed at the final Clinic stage of follow-up where overall attrition differed little between the levels of rurality. The patterns of attrition observed for the entire cohort were almost identical for the subset of 9–15-year-olds used in the later analyses (data not shown).

Table 3: Estimates of the baseline association of BMI with self-rated fitness made using the data available at baseline and each stage of follow-up

	ASHFS (n = 6,398)		**Tracing (n = 5,195)**		**Enrolment (n = 3,970)**		**Questionnaire (n = 3,018)**		**Clinic (n = 1,898)**	
	β	**(SE)***	β	**(SE)***	β	**(SE)***	β	**(SE)***	β	**(SE)***
Self-rated fitness										
Not as fit as most	Ref.		Ref.		Ref.		Ref.		Ref.	
About average fitness	−1.728	(0.099)	−1.789	(0.112)	−1.438	(0.125)	−1.541	(0.142)	−1.454	(0.181)
Fitter than most	−2.325	(0.115)	−2.400	(0.129)	−2.029	(0.143)	−2.107	(0.162)	−1.854	(0.203)
Age	0.629	(0.016)	0.629	(0.018)	0.641	(0.019)	0.650	(0.022)	0.649	(0.026)
Female	0.188	(0.064)	0.162	(0.070)	0.248	(0.077)	0.205	(0.088)	0.132	(0.107)
Constant	12.82	(0.216)	12.91	(0.241)	12.30	(0.267)	12.23	(0.307)	12.14	(0.373)

Note: * β (SE) = regression coefficient (standard error).

Table 4: Demonstration of the use of inverse probability weighting (IPW) and multiple imputation (MI) with the objective of correcting bias in the estimates due to missing data

	ASHFS† (true associations)		Clinic					
			Biased by missing data†		IPW		MI	
	β	(SE)*	β	(SE)*	β	(SE)*	β	(SE)*
Self-rated fitness								
Not as fit as most	Ref.		Ref.		Ref.		Ref.	
About average fitness	−1.728	(0.099)	−1.454	(0.181)	−1.458	(0.244)	−1.731	(0.142)
Fitter than most	−2.325	(0.115)	−1.854	(0.203)	−1.849	(0.255)	−2.231	(0.158)
Age	0.629	(0.016)	0.649	(0.026)	0.649	(0.026)	0.633	(0.019)
Female	0.188	(0.064)	0.132	(0.107)	0.128	(0.106)	0.260	(0.076)
Constant	12.82	(0.216)	12.14	(0.373)	12.14	(0.401)	12.73	(0.272)

Notes: * β (SE) = regression coefficient (standard error).
† These results are repeated from Table 3 to aid interpretation.

The greatest divergence in attrition occurred across the levels of TRSA and SEP (Figure 2). Notably the divergence between high and low levels of TRSA increased through the stages prior to the Clinic stage, while the attrition difference between low and high SEP following Tracing was maintained through the other stages.

Table 3 shows the regression coefficients for the levels of self-rated fitness in the regression of BMI estimated using all baseline data and when estimated using the data available for participants at each stage of follow-up. At baseline those who rated their fitness 'about average' had mean BMI that was 1.79 kg/m^2 lower than those that rated themselves 'not as fit as most'. Those who rated themselves 'fitter than most' had mean BMI that was 2.33 kg/m^2 lower than those who rated themselves 'not as fit as most'. When these coefficients were estimated using the data for those who participated at the Tracing stage, the differences in BMI between the categories of self-rated fitness were unchanged or slightly amplified. The differences in these coefficients were progressively attenuated (biased towards the null) when estimated by limiting the baseline data to participants responding at the Enrolment, Questionnaire and Clinic stages, with the greatest bias observed at the Clinic stage. The bias at the Clinic stage was an attenuation of the magnitude of the BMI difference of 16% for those who rated their fitness 'about average', and of 20% for those who rated their fitness as 'fitter than most', compared with those who rated their fitness 'not as fit as most'.

The estimates for the Clinic stage were used to examine the performance of missing data methods (Table 4). Bias in the model estimates for self-rated fitness were not eliminated using IPW. After correction by IPW, the difference in mean BMI between categories of self-rated fitness was still attenuated by 15% for those rating their fitness 'about average', and by 20% for those rating their fitness as 'fitter than most', compared with those rating their fitness 'not as fit as most'. MI performed much better than IPW and removed the bias in self-rated fitness (the corrected values were within 4% of the true differences). MI did not perform well for the coefficient for females which became biased away from the null, the effect under MI was inflated by approximately 40%.

Discussion

Our study examines a cohort enrolled in childhood with a first follow-up 20 years later in adulthood. We show differential patterns of attrition observed during four sequential stages of the follow-up period: participant tracing, study enrolment, questionnaire completion and clinic attendance. Overall, there is a marked drop in participation through these four stages with only 28.4% of the original cohort completing the Clinic stage. Consistently greater attrition at each stage was observed for subjects with lower scholastic ability, public schooling and lower SEP. Rurality and sex were associated with increased attrition at successive stages prior to the last stage transition, where the associations were reversed. An association between BMI and self-rated fitness estimated using all baseline data was progressively attenuated when estimated using the data available for participants at successive stages of follow-up. Our findings highlight the limitations of IPW in correcting bias when insufficient complete variables are available, while showing that lack of completeness in variables is not a barrier for MI.

Factors associated with attrition in our study are broadly consistent with those identified in previous research. Lower SEP, lower education and male sex have been shown consistently to be predictors of attrition in other studies of child and adolescent participants (Siddiqui et al, 1996; Burkam and Lee, 1998; Fröjd et al, 2010; Gustavson et al, 2012; Mostafa and Wiggins, 2015). In our study, TRSA and SEP were strongly associated with attrition. Education and SEP have been found to be strongly associated with attrition in cohort studies of adults (Ahern and Le Brocque, 2005; Young et al, 2006) and studies of children and adolescents (Burkam and Lee, 1998; Mostafa and Wiggins, 2015). In the BCS70 birth cohort study, Mostafa and Wiggins (2015) suggest that a large drop in participation between ages 16 and 26 could be partially attributable to the change in responsibility for consent from the parents of the participants to the participants themselves as young adults. While we are not able to assess the impact of this, it may also be an influential factor in the overall attrition for our study.

Each follow-up stage after the Tracing stage placed increasing demands on participants. The Enrolment stage required only registration of intent and answering a small array of questions. The questionnaires required increased time commitment, and clinics required a willingness to fast, travel to the clinic location and participate in extensive physical tests and psychological assessments. Patterns of attrition within our study show that factors related to the social environment – school type and SEP – had the greatest influence at the Tracing stage of follow-up, with the overall divergence in attrition between levels of these factors occurring at this early stage. Later follow-up stages showed little further divergence over the initial Tracing spread for these factors. In contrast, a factor related to individual characteristics, TRSA, showed continuing divergence in attrition between levels of this factor during all follow-up stages.

In the assessment of the two missing data techniques, inverse probability weighting produced only partial recovery of the estimates biased by attrition. This is consistent with observations made by Seaman and White (2013): unless informative predictors of missingness are completely observed and available for the weighting model, the performance of this technique is limited. The selection of complete predictors for IPW in our study was limited to just five variables, crucially one of the variables highly predictive of attrition – TRSA – could not be included because it was not complete. Conversely, MI performed well in correcting the bias in estimates of the coefficients for self-rated fitness: the flexibility of including many predictors of the missing values in the BMI and self-rated fitness variables is likely to have contributed to the better performance (Seaman and White, 2013). The apparent inflation in the coefficient for females in the MI analysis is an artefact of the modifying effects of age and self-rated fitness (not reported). The scenario we investigated was that of a baseline association between self-rated fitness and BMI estimated using the data available on participants at each stage of follow-up. Even with attrition greater than 70%, the MI technique was able to correct biased estimates. We anticipate that the scenario we have studied may correspond to missing data assessment for associations measured in follow-up.

This study provides the first demonstration of different missing data patterns at each stage of the sequential follow-up of a large representative cohort. The results help to identity factors associated with missingness that can be informative as predictors in the propensity and imputation models. Knowledge of those factors is critical to an informed assessment of whether the missing data are likely to be MAR or MNAR and informs the construction of models of missingness and the design of sensitivity analyses that are required if the missing data are MNAR. That those factors operate differentially at the various stages of follow-up suggests that stratifying on patterns of missingness may be a productive method for handling missing data (Seaman and White, 2013; Doidge, 2018). The demonstration that correction of bias – in the estimates of a baseline association affected by follow-up – was possible using MI but not IPW highlights the need for rich information to be available for predicting the probability of missingness (IPW) or the conditional distributions of missing variables (MI). It showed that the information needed in each case was specific to the task at hand. Additionally, it showed that the missing data were approximately MAR for self-rated fitness given the available variables in the MI model.

There are some limitations of the study. The results may not generalise to associations other than those between BMI and self-rated fitness or to other cohorts with more regular and/or proximal follow-up. A thorough evaluation of the relative merits of IPW was not possible given the paucity of complete data at baseline due to systematic item non-response (for example, the questionnaire was not administered to participants aged 7–8 years in 1985). The demonstration of correction of bias was specific to a baseline association and does not amount to evidence that bias in analyses with follow-up data can be corrected in the same manner. Additionally, this demonstration was focused on a specific association, between self-rated fitness and BMI. It included a simple adjustment for sex and age, yet these additional covariates have more complex interrelations with each other and the main variables of interest. Reporting these interactions would have required a more complex model and a complicated presentation of results. This was beyond the purposes of this manuscript.

In conclusion, attrition from childhood to adulthood in this 20-year cohort study was influenced by the same factors that have been found to operate in other cohort studies, but these factors operated differentially at each stage of follow-up. Biased estimates in a scenario of extreme attrition were able to be corrected using MI, but a lack of informative variables in the circumstances of this cohort is very likely to have hampered the performance of IPW.

Funding

This work is not directly supported by any specific funding body.

Acknowledgements

The authors would like to acknowledge the hard work by Tim Albion, our data manager, who contributed to data preparation required for this manuscript. Also the authors would like to acknowledge the participants and staff of the ASHFS/CDAH studies who contributed their valuable time and the energy in making this study possible.

Conflict of interest

The authors have no conflicts of interest to declare.

References

Ahern, K. and Le Brocque, R. (2005) Methodological issues in the effects of attrition: simple solutions for social scientists, *Field Methods*, 17(1): 53–69.

Badawi, M.A., Eaton, W.W., Myllyluoma, J., Weimer, L.G. and Gallo, J. (1999) Psychopathology and attrition in the Baltimore ECA 15-year follow-up 1981–1996, *Social Psychiatry and Psychiatric Epidemiology*, 34(2): 91–8.

Brick, J.M. (2013) Unit nonresponse and weighting adjustments: a critical review, *Journal of Official Statistics*, 29(3): 329–53.

Burkam, D.T. and Lee, V.E. (1998) Effects of monotone and nonmonotone attrition on parameter estimates in regression models with educational data: demographic effects on achievement, aspirations, and attitudes, *The Journal of Human Resources*, 33(2): 555–74.

Cole, T.J., Flegal, K.M., Nicholls, D. and Jackson, A.A. (2007) Body mass index cut offs to define thinness in children and adolescents: international survey, *British Medical Journal*, 335(7612): 194–7.

Cumming, J.J. and Goldstein, H. (2016) Handling attrition and non-response in longitudinal data with an application to a study of Australian youth, *Longitudinal and Life Course Studies*, 7(1): 53–63.

Doidge, J.C. (2018) Responsiveness-informed multiple imputation and inverse probability-weighting in cohort studies with missing data that are non-monotone or not missing at random, *Statistical Methods in Medical Research*, 27(2): 352–63.

Eerola, M., Huurre, T. and Aro, H. (2005) The problem of attrition in a Finnish longitudinal survey on depression, *European Journal of Epidemiology*, 20(1): 113–20.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998) An analysis of sample attrition in panel data: the Michigan panel study of income dynamics, *The Journal of Human Resources*, 33(2): 251–99.

Fröjd, S.A., Kaltiala-Heino, R. and Marttunen, M.J. (2010) Does problem behaviour affect attrition from a cohort study on adolescent mental health?, *European Journal of Public Health*, 21(3): 306–10.

Greenland, S. (1977) Response and follow-up bias in cohort studies, *American Journal of Epidemiology*, 106(3): 184–7.

Gustavson, K., von Soest, T., Karevold, E. and Røysamb, E. (2012) Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study, *BMC Public Health*, 12(1): art. 918, https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-12-918

Mostafa, T. and Wiggins, R. (2015) The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies, *Longitudinal and Life Course Studies*, 6(2): 131–46.

Olsen, R.J. (2005) The problem of respondent attrition: survey methodology is key, *Monthly Labor Review*, 128(2): 63–70.

Pyke, J.E. (1987) *Australian Health and Fitness Survey 1985*, Parkside: The Australian Council for Health, Physical Education and Recreation Inc.

R Core Team (2019) *R: A Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing, https://www.R-project.org/

Rubin, D.B. (1976) Inference and missing data, *Biometrika*, 63(3): 581–92.

Seaman, S.R. and White, I.R. (2013) Review of inverse probability weighting for dealing with missing data, *Statistical Methods in Medical Research*, 22(3): 278–95.

Siddiqui, O., Flay, B.R. and Hu, F.B. (1996) Factors affecting attrition in a longitudinal smoking prevention study, *Preventive Medicine*, 25(5): 554–60.

United Directory Systems (2018) *Australia on Disc*, https://www.australiaondisc.com/

van Buuren, S. (2012) *Flexible Imputation of Missing Data*, Boca Raton, FL: CRC Press.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011) mice: multivariate imputation by chained equations in R, *Journal of Statistical Software*, 45(3), https://www.jstatsoft.org/article/view/v045i03

Young, A.F., Powers, J.R. and Bell, S.L. (2006) Attrition in longitudinal studies: who do you lose?, *Australian and New Zealand Journal of Public Health*, 30(4): 353–61. doi: