ORIGINAL ARTICLE

TITLE: Erosion of representativeness in a cohort study

AUTHORS: Christodoulou, M.D.[1]*, Brettschneider, J.A.[2], and Steinsaltz, D.[1]

[1] Department of Statistics, University of Oxford, Oxford, U.K.

[2] Statistics Department, University of Warwick, Coventry, U.K.

* For correspondence e-mail: maria.christodoulou@stats.ox.ac.uk

SUMMARY:

Background: The National Child and Development Study (1958 British Birth Cohort) follows the lives of over 17 000 people born in a single week in England, Scotland, and Wales. After initial recruitment, there have been nine sweeps to gather subsequent life-course data and a Biomedical Sweep (between Sweeps 6 and 7) that has been used in genetic studies. Due to its non-selective recruitment, the survey is frequently used as a representative proxy for the British population in demographic, epidemiological, and medical studies. We examine the effect of attrition on representativeness of female fertility and education length.

Methods: We compare numbers and timings of fertility-related events of female cohort members with national estimates. Spline approximation was used to link records with different aggregation intervals. Participants who were present in the Biomedical Sweep were compared to those who were not.

Results: We established that both timings and counts of maternities and terminations in the cohort diverge from the patterns of their contemporaries. For women who participated in the Biomedical Sweep, we noted positive correlations of study continuation with years spent in full time education, and with age at first birth. We determined that women who did not participate in the Biomedical Sweep reported different fertility patterns from those who did.

Conclusions: For female fertility, the 1958 Birth Cohort is an imperfect proxy for the British population, making the description “broadly representative” potentially misleading. Worsening bias due to attrition and misreporting can be identified and quantified through comparisons with national vital statistics.
Keywords: Bias, attrition, representativeness, fertility, age at menarche, age left full time education, birth cohort, 1958 British Birth Cohort, NCDS, dropouts.

KEY MESSAGES:

- Fertility patterns of women reported in later sweeps of the 1958 British Birth Cohort diverge from those of their contemporaries as estimated from national statistics using spline approximation. Both maternities and terminations are underreported.

- Female participants who dropped out earlier follow national maternity trends more closely than those who participated in the study longer, but they still report slightly fewer maternities than national statistics suggest.

- Female participants who persisted through later sweeps experienced first births later and left education at a later age than those who dropped out earlier.

- Although the 1958 British Birth Cohort has been judged as representative of the British population for some research questions, the cohort population presents increasingly biased fertility patterns in female participants over time. Studies related to fertility using data from this cohort may require adjustment.

INTRODUCTION

The 1958 British Birth Cohort (National Child Development Study) was originally designed as a one-off study focusing on factors surrounding perinatal mortality in Great Britain, and including all births occurring in England, Scotland, and Wales during the first week of March 1958 (1). Cohort members have been followed throughout their lives, creating a rich dataset with extensive records on multiple aspects of their lives, such as health, education, and employment. The non-selective inclusion of every child born in Great Britain in one week has made it an attractive starting point for longitudinal studies of a representative cross-section of the British population. Over 17 000 variables have been collected or derived for the 17 415 initial cohort members (and the 925 added in the replenishment sweeps at ages 7, 11, and 16 for children immigrating to Great Britain), documenting every facet of their lives, from gestation through late middle age. Despite dropouts (due to mortality, emigration, or other reasons), more than half of the cohort (9 377 members) persisted to the Biomedical Sweep (age 44), and more than 40% participated in the most recent sweep in 2013. Levels of attrition are illustrated in Figure 1.

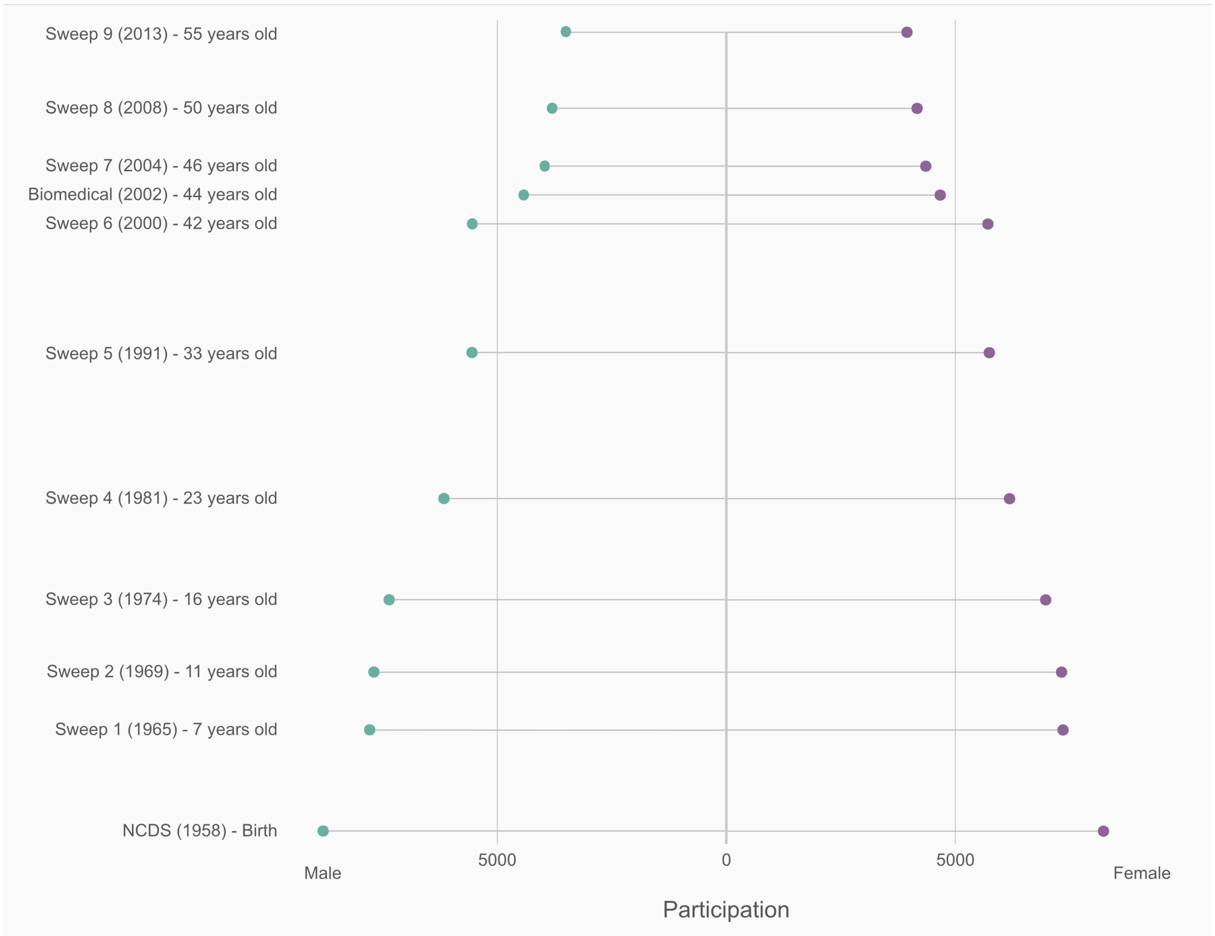

Figure 1: Male and female productive participation by sweep for the 1958 British Birth Cohort (National Child Development Study – NCDS). Starting with 17 415 cohort members at recruitment (8 411 female productive participants and 9 004 male), productive participation declined to 9 377 members for the Biomedical Sweep (4 712 female participants and 4 665 male) in 2002. Age of cohort member and year of Sweep are listed as the start year for each sweep. The highest average annual rate of loss occurred between Sweep 6 and the Biomedical Sweep. This pattern is consistent for both female and male participants. Productive participants are the cohort members who were successfully contacted and responded to a Sweep's questions. As only female participants were released for this study, the male participation was calculated using the publicly available achieved samples for the cohort. The lengths of the lines are proportional to the numbers of individuals they represent.

Because of the outsized role that this cohort has played in a wide range of research, the development of its membership over time has received considerable attention. Atherton et al. (2) compared a variety of sociodemographic, health, and behavioural variables between Sweep 1 and the Biomedical Sweep. They concluded that subject attrition had not substantially vitiated the broadly representative composition of the persisting cohort. This claim has been echoed in numerous publications referring to the cohort as "nationally representative" (3–9) or "broadly representative", subject to attrition (10–16).

The present work examines whether specific data biases may distort the picture of particular research questions within an otherwise representative sample. Our focus is on the impact of attrition on fundamental indicators of education and fertility, such as age at menarche and years spent in education. We explore associations between these variables and the length of continuing participation in the study, for women present in the Biomedical Sweep. We further compare national statistical records to the maternity and termination reports of female participants, separately for those who persisted until the Biomedical Sweep and those who did not. As the Biomedical Sweep was focused on collecting biological material and clinical data, it is only those individuals who were present in this Sweep whose information is routinely released for research that aims to include genetic data. The genetic data commonly used were initially generated by the WTCCC studies (17,18) (Figure 2), which used members of the 1958 Birth Cohort as part of the control group for large meta-analyses on common diseases. These were subsequently made available to other researchers, and have since been used as part of control groups (usually in combination with UK Blood Donors) for a variety of medical studies, including work on leukaemia (19,20), schizophrenia (21) and rheumatoid arthritis (22), as well as for sociological and demographic research (23,24). As these studies include only participants in the Biomedical Sweep, the characteristics of this subgroup are crucial, and it is this that we will take as our primary cutoff.

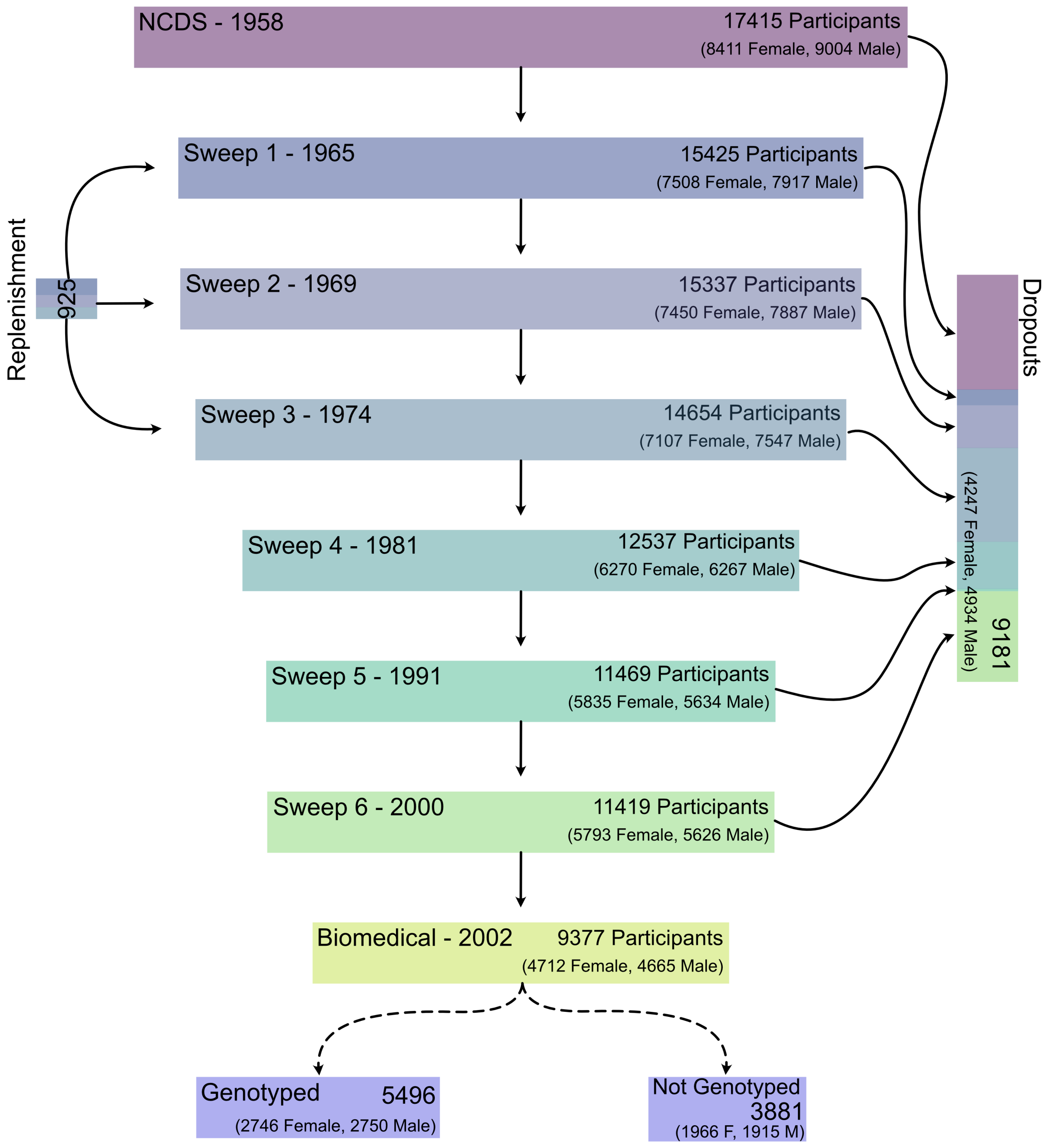

Figure 2: Male and female productive participation, dropouts, and replenishment in the 1958 British Birth Cohort until the Biomedical Sweep, culminating in the numbers of individuals that were genotyped. Productive participants are the cohort members who were successfully contacted and responded to a

sweep's questions. As only female participants were released for this study, the male participation was calculated using the publicly available achieved samples for the cohort. Lengths of rectangles are proportional to the total number of individuals participating.

Whether representativeness as a concept matters for epidemiological, sociological, or demographic research has been extensively debated (25–30), and is beyond the scope of this work. We are not questioning whether overall the 1958 British Birth Cohort is a representative proxy of British population for its age group. Nor are we questioning the appropriateness of the cohort members specifically as a control group for the genetic studies designed by WTCCC. Here we consider the essential yet sometimes neglected distinction between a population being representative overall versus being representative for a particular research objective. We show that the description of the cohort as nationally or broadly representative is potentially misleading for researchers, as it obscures the possible underlying biases that become apparent when framed by a definite research objective.

In this work we demonstrate that annual maternity and termination counts reported by the female participants of the 1958 Birth Cohort diverge from those of their contemporaries, as extracted from the relevant national statistics registries. Specifically, both maternities and terminations of participants appear under-recorded, with those who did not participate in the Biomedical Sweep following national trends more closely than those who did. We note that participants who persisted in the study until the Biomedical Sweep experienced a later age at first birth than those who did not. In this group of women, we further found that the length of continued study participation is positively associated with being older at both first birth and at departure from full-time education.

These findings call into question the suitability of the blanket use of the description of the 1958 Birth Cohort as a "broadly nationally representative cohort". They also highlight the importance of being able to establish both the presence and the impact of bias (31,32), and to take remedial action using qualitative and quantitative methods (33,34).

METHODS

We have instrumentalised the general question of representativeness in two specific questions:

1. whether annual rates of maternity and termination reported by participants who persisted through the later sweeps diverged from rates reported by those who dropped out earlier, and how each of these comports with national statistics. Due to different age aggregation in the latter, we used spline approximation to estimate annual rates, described in the Section: "Annual maternity and termination rate estimates"; and
2. whether participant attrition interacts with education and fertility variables. This was conducted using time-to-event analysis and is presented in the Section: "Attrition through the sweeps".

Combined, these delineate the impact of attrition on the overall makeup of the cohort for fundamental fertility and education variables. It should be noted that miscarriages have not been examined, as no comparable national statistics are available.

DATA AND SOFTWARE

All available fertility-related variables from the 1958 Birth Cohort were requested from METADAC and used to reconstruct in-depth fertility profiles for the 4 712 female participants in the Biomedical Sweep conducted in 2002. We denote by "year of Sweep" the year in which a given sweep started. Protocols for clean-up (the accuracy of which was tested using synthetic data) are supplied in the **Supplementary Material**. Variables relating to participation in sweeps, and age of leaving full-time education were also requested. A list of all 942 requested variables is provided in the **Supplementary Material**. As we were interested in the impact of attrition, we further requested information about the female cohort members who did not participate in the Biomedical Sweep. This yielded 4 247 additional participants, with 314 variables requested (a subset of the original variable request, relating to the available sweeps). No data, if available were released for Sweep 6 for this group. Although we noted some slight discrepancies between the numbers of participants in our data and those recorded on published tables in Atherton et al.(2) and Power et al (1), we are in overall agreement on the productive participant counts. As part of our quality-control assessment, we noted that initial recruitment sex ratio (107.05 males to 100 females) matches the birth sex ratio of the general population. Further consistency checks are detailed in the associated R scripts.

Maternities (live births and stillbirths) and terminations for all women living in Great Britain born in 1958 were obtained from the Office for National Statistics (ONS) and from National Records of Scotland (NRS) (35), for all years between 1972 and 2001. ONS maternity and termination records (covering England and Wales) prior to 1992 were available only as age-group aggregates through the Conception Statistics tables (36,37). From NRS both age-group aggregates and a breakdown by mother's age were available (35). The life table for women born in 1958 was obtained from the Human Mortality Database (38). Statistical analysis was conducted using base R (39), "tidyverse" (40), and "survival" (41) through RStudio (42). All code is available in the **Supplementary Material**.

ANNUAL MATERNITY AND TERMINATION RATE ESTIMATES

We compared rates of maternities and terminations reported by participants with the corresponding rates in the general population, derived from official statistics. Although counts by individual years of mother's age are available from the Scottish data, we aggregated the age groups to align them with the data releases from England and Wales. To then extract single-year estimates of maternity and termination rates for individuals born in 1958 we fitted splines, for which we tested three different specific procedures: a standard b-spline (knots at 1974, 1978, 1983, 1988, and 1993 – corresponding to the years in which participants crossed over between age classes of the vital statistics tables, for example 1974 when they entered the 16-19 years age class – a natural cubic spline (knots as in the b-spline), and a quadratic optimisation spline proposed by Grigoriev et al. (43). This last is based on minimising the sum of squared second-order differences, yielding annual rate estimates (per 1000 women) for maternity and terminations. These were compared with the per-1000-women rates for maternities and terminations within the 1958 Birth Cohort, for which we also calculated 95% confidence intervals.

SPLINE VERIFICATION

We tested the ability of the spline methodology to recover the known individual-year counts for Scottish maternities from multi-year aggregated Scottish maternity data. We transformed the Scottish birth counts into rates per 1000 women by using the life table from the Human Mortality Database. These are plotted together with Scotland-only spline estimates in Figure 3. As the quadratic-programming spline fitted the count rates better than the natural or b-splines, we rely primarily on this method in our subsequent analysis.

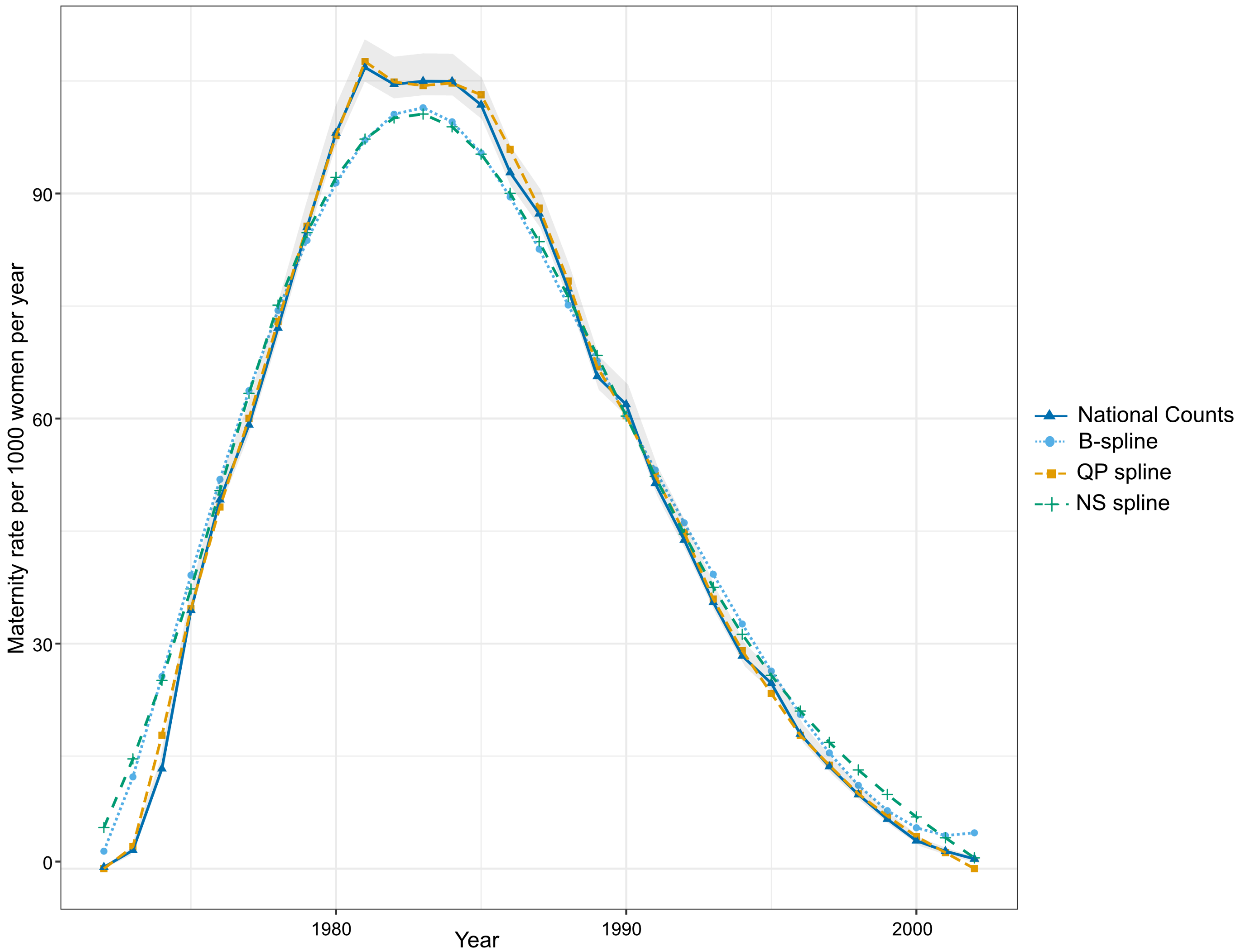

Figure 3: Maternity spline estimates extracted from age group aggregates against single-year counts for Scotland. Three spline methods are presented: B-splines (light blue), Quadratic programming splines (yellow), and Natural cubic splines (green). These are compared with rates per 1000 women estimated from counts (dark blue, with 95% confidence interval in light grey).

ATTRITION THROUGH THE SWEEPS

We distinguished between two broad categories of participation: participating in the Biomedical Sweep ("Persisters") or not ("Dropouts"). We define two different measures of participation: the last sweep of having participated in any way ("Overall Continuation Length") and the last sweep of having responded to fertility-related questions ("Fertility Continuation Length").

PARTICIPATION INDICATORS	**Persister:** cohort member who participated in the Biomedical Sweep.
	Dropout: cohort member who did not participate in the Biomedical Sweep.
MEASURES OF CONTINUATION LENGTH	**Overall Continuation Length:** last sweep in which cohort member has participated in any way
	Fertility Continuation Length: last sweep in which cohort member has responded to any fertility-related questions

Table 1: Glossary of terminology used, defining participation indicators (Persisters vs Dropouts) and measures of Continuation Length (Overall vs Fertility).

WITHIN THE PERSISTERS

We examined associations between fertility and education event times: *age at menarche*, *age at first birth*, and *age left full time education*. Age at menopause was not studied as most of the sweeps occurred prior to the menopause age range. Kendall's tau was used as a correlation measure between ages at dropout and at event. Analyses were conducted using both Continuation Length measures.

For individuals lacking event times, censoring time was determined based on auxiliary information present in the dataset. For example, for *age at first birth*, individuals with no recorded maternities were considered censored at the last time when a birth could have been observed using other fertility variables where available (such as menopause age). Where insufficient information was available to make a clear judgement about censoring time, these individuals were excluded entirely from the analysis. The numbers of excluded Persisters are summarised in Table 2.

	Age at menarche		*Age at first birth*		*Age left full time education*	
	Excluded	Included	Excluded	Included	Excluded	Included
Overall Continuation Length	0	4712	414	4298	239	4473
Fertility Continuation Length	36	4676	438	4274	275	4437

Table 2: Numbers of Persisters that were included and excluded in the analyses for *age at menarche*, *age at first birth* and *age left full time education*.

PERSISTERS VS DROPOUTS

We compared *age at menarche* and *age at first birth* between the 4 712 Persisters and the 4 247 Dropouts using log-rank tests. *Age left full time education* was not studied here as the relevant variables were only collected in Sweep 6, when many Dropouts had already left the study. Age at menopause was not studied because the last records for Dropouts would be prior to age 44, which is before most women would have experienced menopause.

RESULTS

ANNUAL MATERNITY AND TERMINATION RATE ESTIMATES

Annual maternity rates for Persisters and Dropouts in the Birth Cohort, as well as estimates of these rates from national statistics, are shown in Figure 4. Until 1987, both Persisters and Dropouts show lower rates than national statistics, but Dropouts follow national statistics more closely than Persisters. Fewer maternities were recorded for Persisters between 1973 and 1984 than would be predicted from the records of the general population, with an average of 2.21% fewer maternities per year (sd: 0.73%, range: 1.01% - 3.13%). In 1982 and 1983 the deviation is particularly large, and shows a prominent dip in Persisters' maternity rate curve, separating them even more from Dropouts. From 1988 onwards Persisters' maternity rates exceed the rates calculated from national statistics.

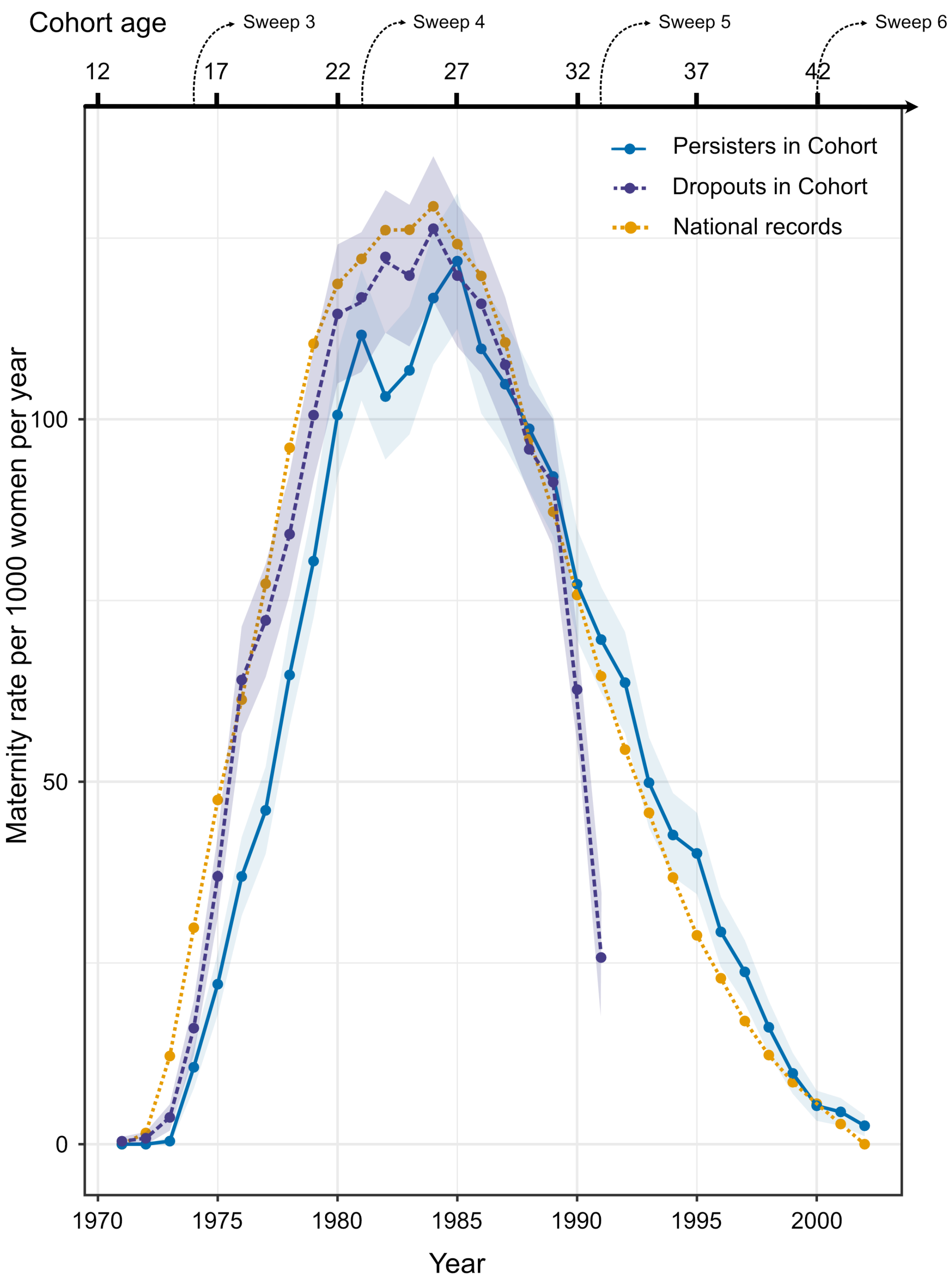

Figure 4: Annual maternity rates (per 1000 women) for Persisters (dark blue, with 95% confidence interval in lighter shade) and Dropouts (purple, with 95% confidence interval in lighter shade) in the Birth

Cohort. These are compared to quadratic programming spline approximations of these rates based on national statistics age group aggregates (yellow).

Maternity rate differences between the cohort participants and National statistics are presented in Figure 5.

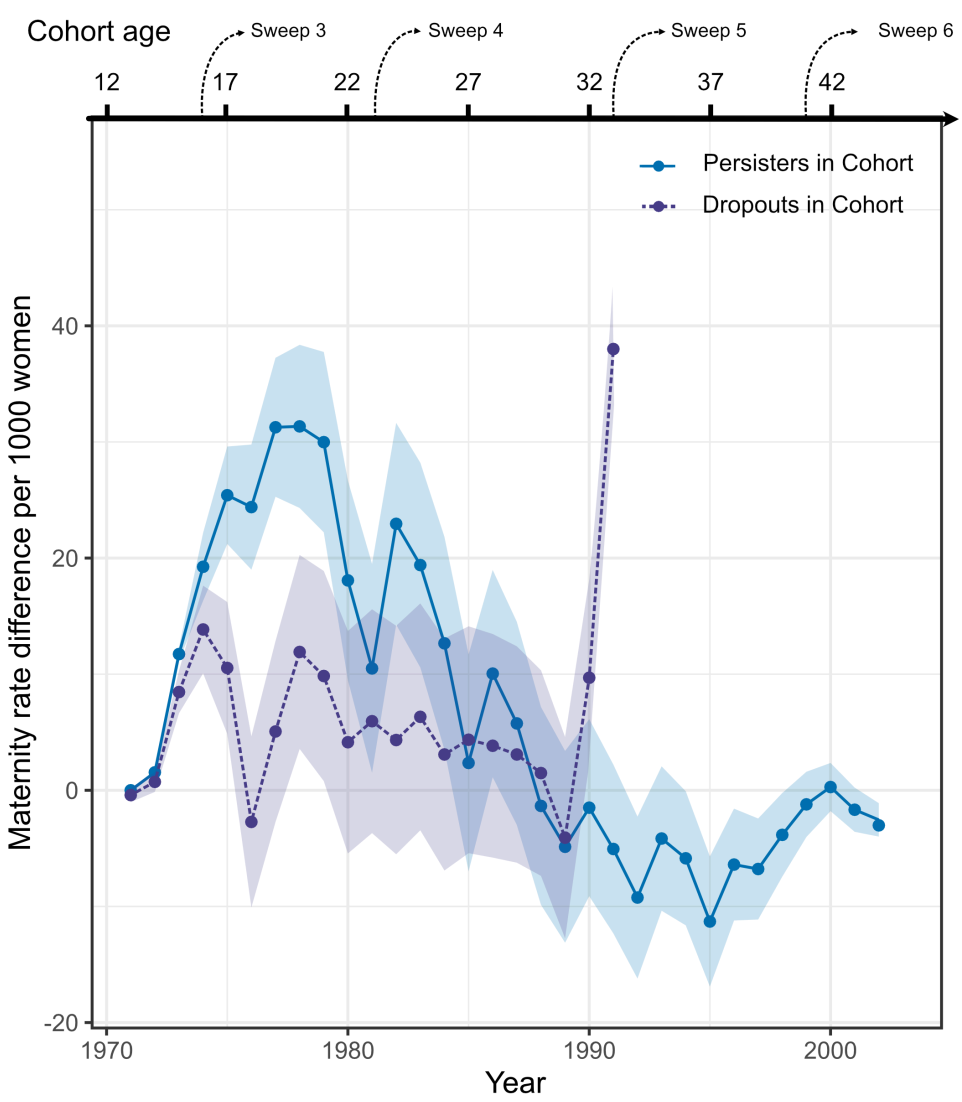

Figure 5: Annual maternity rate difference (per 1000 women) between National statistics and Persisters (dark blue, with 95% confidence interval in lighter shade), and National statistics and Dropouts (purple, with 95% confidence interval in lighter shade). For National statistics, rates were estimated using quadratic programming spline approximations of age group aggregates.

Comparison of termination rates between national statistics and those reported by Persisters is illustrated in Figure 6. The reported termination rates for Persisters diverge substantially from those approximated from the national statistics data. The terminations for Dropouts have not been included as the events were not linked with specific years. This is due to the manner termination-related questions were asked prior to 1990s.

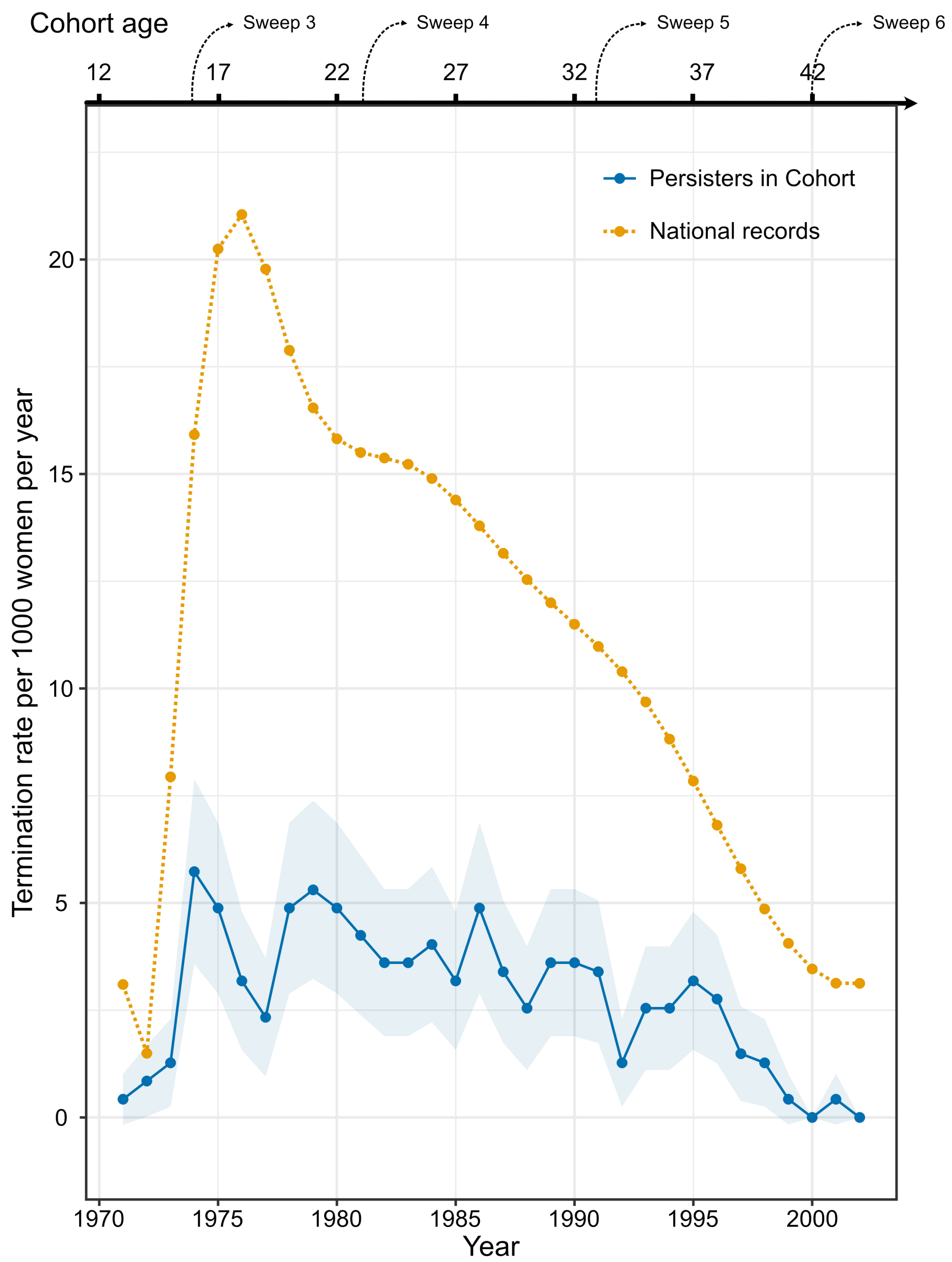

Figure 6: Annual termination rates (per 1000 women) as recorded for Persisters in the Birth Cohort (dark blue, with 95% confidence interval in lighter shade) compared to quadratic programming spline approximations of these rates based on national statistics age group aggregates (yellow).

ATTRITION THROUGH THE SWEEPS

WITHIN PERSISTERS

Kendall's tau between Continuation Length past the Biomedical Sweep and three event times – *age at menarche*, *age at first birth*, and *age left full time education* – are summarised in Table 3. If measured by "Overall Continuation Length", the length of continuation is correlated with each of them – weakly negatively for *age at menarche*, positively for *age at first birth*, and *age left full time education*. In other words, members who participated in later sweeps experienced menarche at a younger age, while they were older when they experienced their first birth and when they left full time education. If measured by "Fertility Continuation Length", the negative correlation of Continuation Length with *age at menarche* disappears, but the positive correlations with the other two event times remain.

	Age at menarche	*Age at first birth*	*Age left full time education*
Overall Continuation Length	τ= -0.033 95% CI: (-0.045, -0.021) p-value=0.009	τ=0.124 95% CI: (0.112, 0.136) p-value<2.2×10^{-16}	τ=0.117 95% CI: (0.106, 0.127) p-value<2.2×10^{-16}
Fertility Continuation Length	τ= -0.003 95% CI: (-0.02, 0.013) p-value=0.778	τ=0.137 95% CI: (0.12, 0.154) p-value<2.2×10^{-16}	τ=0.138 95% CI: (0.122, 0.153) p-value<2.2×10^{-16}

Table 3: Kendall's tau rank correlation between Continuation Length and three event times: *age at menarche*, *age at first birth*, and *age left full time education.* τ estimates, their 95% confidence intervals, and the associated p-values were calculated for two participation indicators (Overall Continuation Length and Fertility Continuation Length) resulting in similar results for *age at first birth* and *age left full time education*, but different results for *age at menarche*.

PERSISTERS VS DROPOUTS

Results of the comparison for *age at first birth* between Persisters and Dropouts in the Birth Cohort are presented in Figure 7. Dropouts experienced an earlier age at first birth, with median age 22, as opposed to 25 for Persisters.

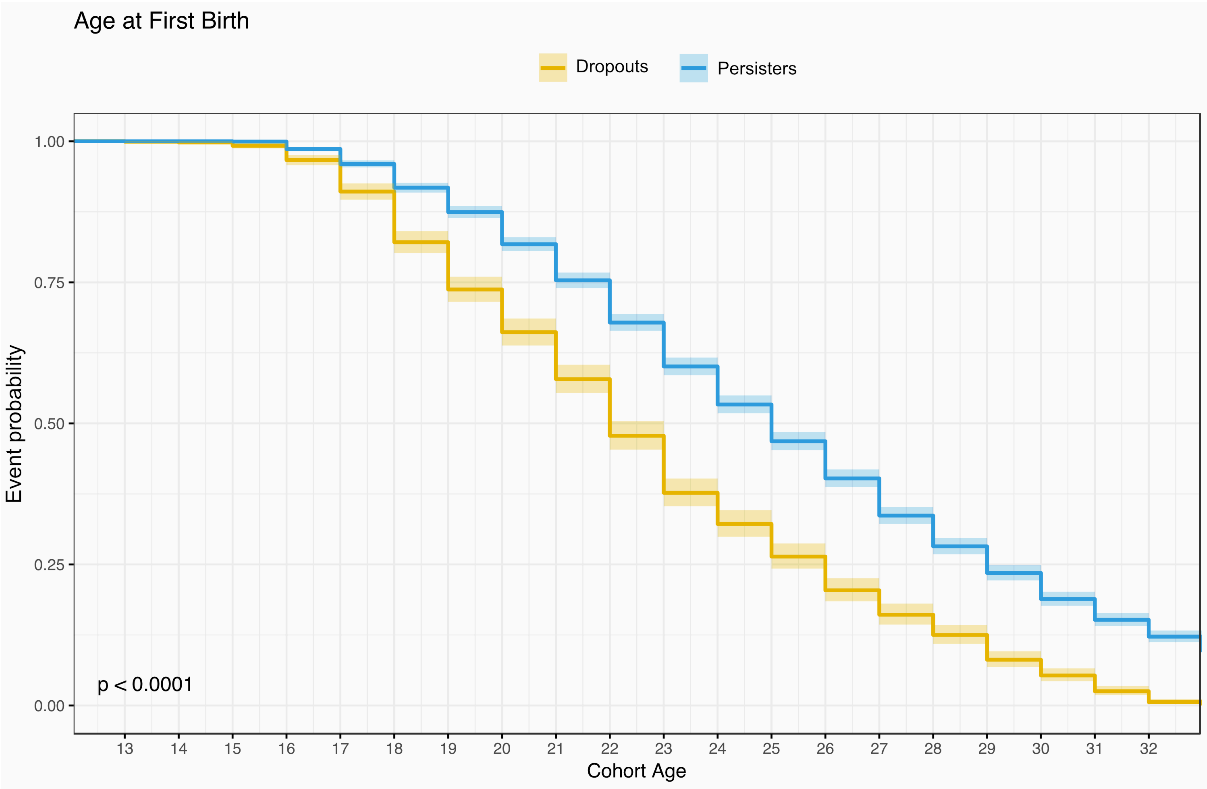

Figure 7: Kaplan-Meier curves for *age at first birth* comparison between Dropouts in the Birth Cohort (yellow) and Persisters in the Birth Cohort (in blue).

Comparison of *age at menarche* between Persisters and Dropouts in the Birth Cohort is presented in Figure 8. The two curves are similar, although proportionally more participants in the group of Persisters experience menarche at age 13.

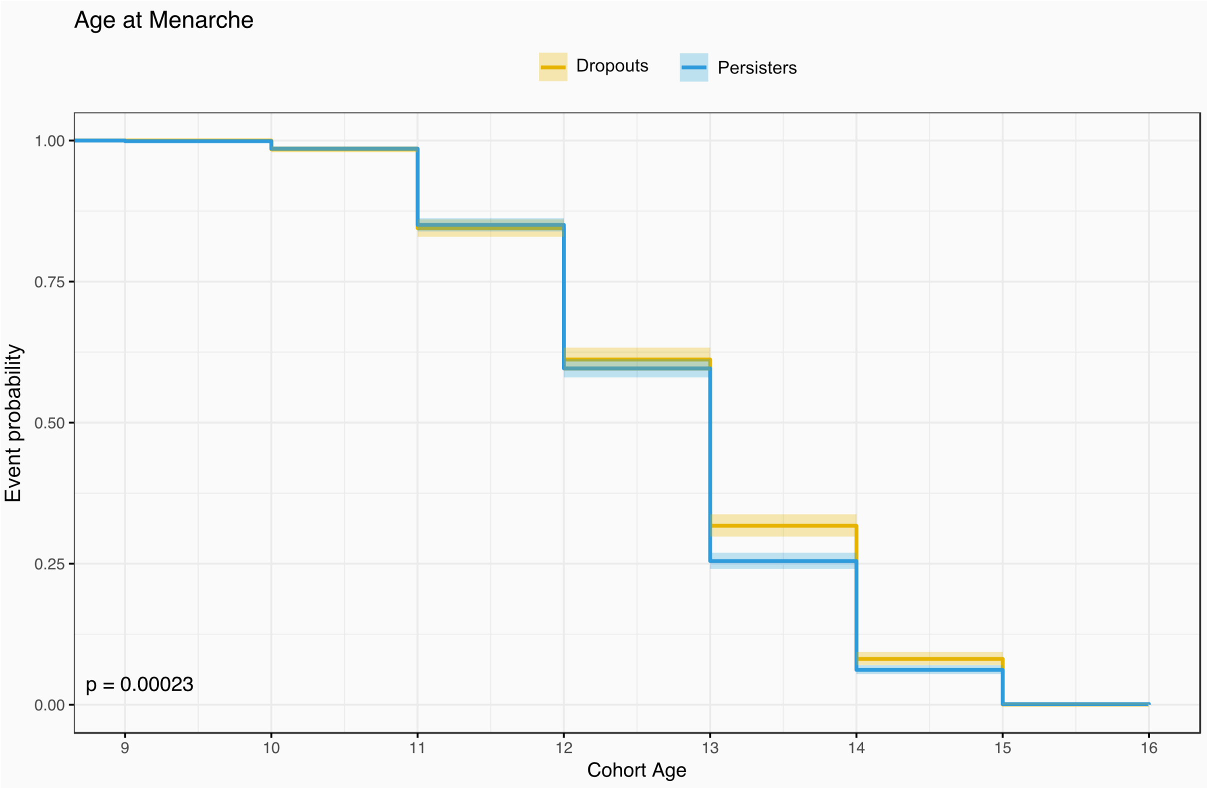

Figure 8: Kaplan-Meier curves for *age at menarche* comparison between Dropouts in the Birth Cohort (yellow) and Persisters in the Birth Cohort (blue).

DISCUSSION

We have established two key findings about the impact of attrition on basic fertility and education characteristics for the women of the 1958 Birth Cohort. Our distinction between Persisters (women present in the Biomedical Sweep) and Dropouts (women who did not participate in the Biomedical Sweep) seems to be a productive division for understanding differential continuation rates. Suitable population estimates are often unobtainable, making direct comparisons between a study cohort and the target population impossible (44). We confirmed the utility of the quadratic-optimisation splines for calculating annual rates for events where only age aggregates were archived.

1. Fewer fertility events are reported in the cohort than would be expected from national statistics. For maternities this discrepancy lingers into the mid 1980s for Persisters and, to a lesser degree, also for Dropouts. Reported terminations for Persisters are extremely low throughout, as compared with the known rates from national statistics.
2. For Persisters both *age at first birth* and *age left full time education* are positively correlated with their length of continued study participation. Persisters experienced their first birth at a later age than Dropouts. *Age at menarche* was similar overall between the two groups.

These findings suggest that Persisters have different fertility and educational profiles to Dropouts. Further, this profile diverges from national trends of their contemporaries for both maternities and terminations. Because of the nature of its initial recruitment, the 1958 Birth Cohort is often considered a representative sample of the British population. Our findings suggest that this perception is not entirely accurate, at least with regard to the fertility events considered here. Following the *Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement* (45) we recommend that this disparity be mentioned in future work relying on the representativeness of the cohort for fertility questions.

By comparing the maternity patterns of Persisters with National Statistics we established that there were fewer maternities reported by the cohort participants between 1973 and 1984 (Figure 4). The disparity between the 1958 Birth Cohort fertility patterns and the estimates based on national statistics is perplexing. We had expected self-reported births to yield numbers and age patterns very close to those registered by statistics agencies. The discrepancy suggests either a reliability issue with the maternity records for Persisters in the Birth Cohort – underreporting or recording errors – or a genuine divergence in behaviour between Persisters and the general population. While less pronounced, Dropouts also exhibit lower maternities rates than national statistics suggest. Explanations for the observed differences need to address both the smaller numbers overall and the additional gaps due to non-random attrition. Of some interest is also the slight decrease in rates for Persisters in 1982 and 1984. These are the years following Sweep 4, making it possible for participants to have mistakenly assumed during Sweep 5 that they had already reported a particular birth. This explanation does not differentiate between Dropouts and Persisters, however. As this particular dip is not observed in the Dropouts, it seems likely that there is an artefact of the data collection that affects predominantly the Persisters.

Terminations appear to be severely underreported for Persisters (Figure 6). Underreporting of terminations is potentially consequential for reproductive and medical studies (46). For example, biased termination records have led to erroneous reports of association between terminations and breast cancer, which were only later disproved through laborious large-scale studies (47,48). The underreporting is likely due to social-desirability bias. Specifically, there are two aspects of study design for the 1958 Birth Cohort that may have exacerbated it. The first is that most of the fertility data were elicited in interviews, whether face-to-face or by telephone, rather than in self-administered questionnaires which might mitigate social-desirability bias. The second is the change in fertility questionnaire structure between Sweep 4 (in 1981) and Sweep 5 (in 1991). Whereas the participants in Sweep 4 were asked to list any natural children they may have, in birth order, from Sweep 5 onwards participants were asked specifically about their pregnancies, starting from the most recent and going backward in time. Miscarriages and terminations from Sweep 4 are reported only collectively, as total numbers, without associated dates. National statistics (Figure 6) suggest that the terminations for women born in 1958 peaked by 1976. This means that the increased resolution in data collection for the cohort, introduced in Sweep 5, appears too late. The responses to the pregnancy questions from 1991 onwards, although informative in assisting to reconstruct the past, may be compromised by recall bias on top of the inevitable social-desirability bias, and the aforementioned non-random attrition.

For Persisters, we have found that participation in the final four sweeps (Biomedical and Sweeps 7-9) was associated with being older both at first birth and at departure from full-time education. That these two event times follow similar patterns might have been anticipated from positive correlation established in the literature (49,50), but we know of no reason to have predicted an association between the timings of those events and Continuation Length in the study.

The correlation between Continuation Length and the above-mentioned two event times becomes stronger when continuation is defined by a subject's engagement specifically with the fertility questions. How to measure Continuation Length is an important consideration, especially within the context of any statistical analyses of fertility – for instance, comparing menopause timing – that try to account for right censoring of missing subjects. Using Overall Continuation Length as the censoring time for individuals who lack event times may be misleading if the individuals have not interacted with any of the fertility questions. Fundamentally, the last time when we can confidently assume that an event has not occurred is the last time when we have been explicitly or implicitly told it has not occurred. When subjects have only responded to some questions we may be left unsure whether they are, in fact, censored.

Whether overall the 1958 British Birth Cohort is a good proxy for the British population of this age group – in other words, whether it is representative of the population – is a multifaceted question that defies any simple answer. The general definition of representativeness of cohorts in research on human populations, and whether or not it is an essential quality of a study, has been discussed extensively (25–30). The analyses presented here illustrate some of the ways in which representativeness may depend on the research question under consideration. This has implications for the use of the cohort study data. Once a research question has been defined, relevant biases can be established and, in some cases, quantified, for example by means of causal diagrams, relative odds ratios, or expert assessment (31,32). Then, depending on the combined nature of the bias and the research question, remedial action can be considered – whether it is a quantitative correction such as multiple imputation with inverse probability weighting (34,51), or by other means, such as those suggested in (33). Since it is *a priori* not known if the effect of biases in associations may be very modest, as in (31), or quite substantial, as in (34), it is important to determine how a specific research question may interact with the identified biases.

ETHICS

Access to anonymised linked records from NCDS was obtained after approval by the METADAC panel on April 11th 2017 (application number: MDAC-2017-0007-01).

FUNDING

This work was supported by the Biotechnology and Biological Sciences Research Council [BB/S001824/1]

ACKNOWLEDGMENTS

Data governance was provided by the METADAC data access committee, funded by ESRC, Wellcome, and MRC. (Grant Number: MR/N01104X/1)

This work made use of data and samples generated by the 1958 Birth Cohort (NCDS), which is managed by the Centre for Longitudinal Studies at the UCL Institute of Education, funded by the Economic and Social Research Council (grant number ES/M001660/1). Access to these resources was enabled via the Wellcome Trust & MRC: 58FORWARDS grant [108439/Z/15/Z] (The 1958 Birth Cohort: Fostering new Opportunities for Research via Wider Access to Reliable Data and Samples). Before 2015 biomedical resources were maintained under the Wellcome Trust and Medical Research Council 58READIE Project (grant numbers WT095219MA and G1001799).

REFERENCES

1. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). Int J Epidemiol. 2006 Feb 1;35(1):34–41.
2. Atherton K, Fuller E, Shepherd P, Strachan DP, Power C. Loss and representativeness in a biomedical survey at age 45 years: 1958 British birth cohort. J Epidemiol Community Health. 2008;62(3):216–23.
3. Capon F, Di Meglio P, Szaub J, Prescott NJ, Dunster C, Baumber L, et al. Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis. Hum Genet. 2007;122(2):201–6.
4. Simpson CL, Hysi P, Bhattacharya SS, Hammond CJ, Webster A, Peckham CS, et al. The roles of PAX6 and SOX2 in myopia: Lessons from the 1958 British Birth Cohort. Investig Ophthalmol Vis Sci. 2007;48(10):4421–5.
5. Capon F, Bijlmakers M-J, Wolf N, Quaranta M, Huffmeier U, Allen M, et al. Identification of ZNF313 / RNF114 as a novel psoriasis susceptibility gene. Hum Mol Genet. 2008 Jul 1;17(13):1938–45.
6. Razavi T, Clark C, Stansfeld SA. Work-family conflict as a predictor of common mental disorders in the 1958 British birth cohort. Longit Life Course Stud. 2015;6(3):264–78.
7. Kwok MK, Yeung SLA, Leung GM, Mary Schooling C. Birth weight and adult cardiovascular risk factors using multiple birth status as an instrumental variable in the 1958 British Birth Cohort. Prev Med (Baltim). 2016;84:69–75.
8. Staff J, Maggs J. Alcohol and Cigarette Use From Ages 23 to 55: Links With Health and Well-Being in the Long-Term National Child Development Study. J Stud Alcohol Drugs. 2017;78(3):394–403.
9. Evans-Lacko S, Takizawa R, Brimblecombe N, King D, Knapp M, Maughan B, et al. Childhood bullying victimization is associated with use of mental health services over five decades: A longitudinal nationally representative cohort study. Psychol Med. 2017;47(1):127–35.
10. Cooper R, Atherton K, Power C. Gestational age and risk factors for cardiovascular disease: Evidence from the 1958 British birth cohort followed to mid-life. Int J Epidemiol. 2009;38(1):235–

44.

11. Vimaleswaran KS, Cavadino A, Hyppönen E. Evidence for a genetic interaction in allergy-related responsiveness to vitamin D deficiency. Allergy. 2012 Aug;67(8):1033–40.

12. Vimaleswaran KS, Cavadino A, Hyppönen E. APOA5 genotype influences the association between 25-hydroxyvitamin D and high density lipoprotein cholesterol. Atherosclerosis. 2013 May;228(1):188–92.

13. Pinto Pereira SM, Li L, Power C. Early-life predictors of leisure-time physical inactivity in midadulthood: Findings from a prospective british birth cohort. Am J Epidemiol. 2014;180(11):1098–108.

14. Krzyzanowska M, Mascie-Taylor CGN, Thalabard JC. Is human mating for height associated with fertility? Results from a british national cohort study. Am J Hum Biol. 2015;27(4):553–63.

15. Lacey RE, Sacker A, Kumari M, Worts D, McDonough P, Booker C, et al. Work-family life courses and markers of stress and inflammation in mid-life: Evidence from the National Child Development Study. Int J Epidemiol. 2016;45(4):1247–59.

16. Cherrie MPC, Sarran C, Osborne NJ. Association between serum 25-hydroxy vitamin D levels and the prevalence of adult-onset asthma. Int J Environ Res Public Health. 2018;15(6):1–12.

17. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet. 2007;39(11):1329–37.

18. Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661–78.

19. Speedy HE, Di Bernardo MC, Sava GP, Dyer MJS, Holroyd A, Wang Y, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2014;46(1):56–60.

20. Law PJ, Sud A, Mitchell JS, Henrion M, Orlando G, Lenive O, et al. Genome-wide association analysis of chronic lymphocytic leukaemia, Hodgkin lymphoma and multiple myeloma identifies pleiotropic risk loci. Sci Rep. 2017;7(January):1–11.

21. Richards AL, Leonenko G, Walters JT, Kavanagh DH, Rees EG, Evans A, et al. Exome arrays

capture polygenic rare variant contributions to schizophrenia. Hum Mol Genet. 2016 Mar 1;25(5):1001–7.

22. Vignal CM, Bansal AT, Balding DJ. Using Penalised Logistic Regression to Fine Map HLA Variants for Rheumatoid Arthritis. Ann Hum Genet. 2011 Nov;75(6):655–64.

23. Bukodi E, Goldthorpe JH. Educational attainment - relative or absolute - as a mediator of intergenerational class mobility in Britain. Res Soc Stratif Mobil. 2016;43:5–15.

24. Bukodi E, Goldthorpe JH, Waller L, Kuha J. The mobility problem in Britain: New findings from the analysis of birth cohort data. Br J Sociol. 2015;66(1):93–117.

25. Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. Int J Epidemiol. 2013;42(4):1012–4.

26. Elwood JM. Commentary: On representativeness. Int J Epidemiol. 2013;42(4):1014–5.

27. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years-has the time come to move on? Int J Epidemiol. 2013;42(4):1016–7.

28. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. Int J Epidemiol. 2013;42(4):1018–22.

29. Ebrahim S, Smith GD. Commentary: Should we always deliberately be non-representative? Int J Epidemiol. 2013;42(4):1022–6.

30. Rothman KJ, Gallacher JEJ, Hatch EE. Rebuttal: When it comes to scientific inference, sometimes a cigar is just a cigar. Int J Epidemiol. 2013;42(4):1026–8.

31. Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. Acta Obstet Gynecol Scand. 2018;97(4):407–16.

32. Thompson S, Ekelund U, Jebb S, Lindroos AK, Mander A, Sharp S, et al. A proposed method of bias adjustment for meta-analyses of published observational studies. Int J Epidemiol. 2011;40(3):765–77.

33. Ioannidis JPA. Commentary: Adjusting for bias: A user's guide to performing plastic surgery on meta-analyses of observational studies. Int J Epidemiol. 2011;40(3):777–9.

34. Munafò MR, Tilling K, Taylor AE, Evans DM, Smith GD. Collider scope: When selection bias can substantially influence observed associations. Int J Epidemiol. 2018;47(1):226–35.

35. National Records of Scotland. Births Time Series Data - Births, by mother's age, Scotland, 1945

to 2018. 2019. p. BT2.

36. Office for National Statistics. Dataset: Conception statistics. 1985.

37. Office for National Statistics. Dataset: Conception statistics [Internet]. 2003 [cited 2018 Jan 22].

38. University of California Berkeley (USA), Max Planck Institute for Demographic Research (Germany). Human Mortality Database [Internet]. [cited 2018 Aug 18].

39. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. [Internet]. Vienna, Austria.: R Foundation for Statistical Computing; 2019.

40. Wickham H. tidyverse: Easily Install and Load the "Tidyverse". [Internet]. 2017.

41. Therneau T. A Package for Survival Analysis in S version 2.38 URL: https://CRAN.R-project.org/package=survival. 2015.

42. R Studio Team. RStudio: Integrated Development for R [Internet]. Boston, MA; 2018.

43. Grigoriev P, Jdanov DA, Shkolnikov VM. New methods for estimating detailed fertility schedules from abridged data. 2018;49(0):0–27.

44. Brilleman SL, Pachana NA, Dobson AJ. The impact of attrition on the representativeness of cohort studies of older people. BMC Med Res Methodol. 2010;10.

45. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. Epidemiology. 2007;18(6):800–4.

46. Fu H, Darroch JE, Henshaw SK, Kolb E. Measuring the extent of abortion underreporting in the 1995 National Survey of Family Growth. Fam Plann Perspect. 1995;30(3):128–33, 138.

47. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and abortion: Collaborative reanalysis of data from 53 epidemiological studies, including 83 000 women with breast cancer from 16 countries. Lancet. 2004;363(9414):1007–16.

48. Stuart GS, Grimes DA. Social desirability bias in family planning studies: a neglected problem. Contraception. 2009;80(2):108–12.

49. Rendall M, Aracil E, Bagavos C, Couet C, DeRose A, DiGiulio P, et al. Increasingly heterogeneous ages at first birth by education in Southern European and Anglo-American family-policy regimes: A seven-country comparison by birth cohort. Popul Stud (NY). 2010;64(3):209–

27.

50. Bongaarts J, Mensch BS, Blanc AK. Trends in the age at reproductive transitions in the developing world: The role of education. Popul Stud (NY). 2017;71(2):139–54.

51. Mostafa T, Wiggins RD. The impact of attrition and non-response in birth cohort studies: A need to incorporate missingness strategies. Longit Life Course Stud. 2015;6(2):131–46.