

# **THE ALSPAC STUDY**

## **MZ FILE**

### **SAMPLE DATA COLLECTED FROM A VARIETY OF SOURCES**

**Prepared by**

**The ALSPAC Study Team**

**Documentation giving frequencies, background and instructions for use.**

**Last updated for version 5a of the file**

**January 2009**

**© University of Bristol**

# Contents

Page 3	Introduction
Page 5	Membership of core sample
Page 5	Mothers with more than one pregnancy in core sample
Page 6	Pregnancy size
Page 7	Pregnancy outcome pre 20 weeks
Page 8	Survival
Page 9	Season of conception
Page 10	Gestation at enrolment
Page 12	Gestation at delivery
Page 15	Season of birth
Page 16	Maternal age
Page 21	Sub-studies

## Introduction

### Background

The MZ file contains basic data on the women whose pregnancies are in the ALSPAC sample. The variables include: number of eligible pregnancies, pregnancy size and outcome, survival of children, seasons of conception and birth, gestation at enrolment and delivery and maternal age.

### Study eligibility criteria

To be eligible for the study, women had to be resident in that part of the County of Avon that was in the South West Regional Health Authority while pregnant. In addition, their expected date of delivery had to lie between 1<sup>st</sup> April 1991 and 31<sup>st</sup> December 1992.

### Eligible cohort

Ideally the MZ file would contain a record for every pregnancy that met the study eligibility criteria. However, it was not possible to collate such a list during the recruitment phase of the study in the early 1990s. Identification of the whole eligible cohort was first attempted in around 1998 using hospital records. However, the 1998 identification work is currently being reviewed. It is hoped that this work will be completed during 2009 and when it is a new version of the MZ file will be created that contains every eligible pregnancy. Note that it will not be possible to exactly identify the whole eligible cohort. In particular, it will not be possible to identify eligible pregnancies that ended in early miscarriage.

### Core ALSPAC sample

The core ALSPAC sample consists of **14,541** pregnancies. This was defined by version 4 of the MZ file as the number of eligible pregnancies for which the mother enrolled in the ALSPAC study and had either returned at least one questionnaire or attended a “Children in Focus” clinic by 19/07/99. The measure of whether or not a mother had returned any questionnaires was obtained by looking at the A-J, MA, PA-PE, KA-KF and KJ built files as they stood at that time.

Prior to the definition of the core ALSPAC sample on 19/07/99 various larger numbers were used at different times to refer to the size of the cohort. As well as including some pregnancies for which the mother enrolled but never returned any data these older numbers also included some mothers that enrolled twice for the same pregnancy and some pregnancies were not actually eligible (e.g. the mother lived outside of the defined geographical area).

### Non-core cases

Eligible cases not in the core sample are referred to as “non-core cases”. There are two main mechanisms by which data are available for non-core cases:

- The main source of information on non-core cases is linkage to external records. For reasons of confidentiality ALSPAC is not allowed to divulge membership of the study. However, membership of the eligible cohort is regarded as public domain information. Linkage to external records is therefore always done on the basis of the eligible cohort.
- When the oldest children were approximately 7 years of age, an attempt was made to bolster the core sample with eligible cases that failed to join the study originally. ALSPAC collected data (e.g. questionnaires, Focus visits) are available for about 550 eligible cases not in the core sample.

### **Sample on current version of the MZ file**

The sample on the current version of the MZ file consists of the 14,541 pregnancies in the core ALSPAC sample plus 5,247 non-core cases for which data appear on the current set of published data files.

The total number of cases on this file should not be thought of as a figure of any significance. In particular, care should be taken not to imply that it is the total number of eligible cases. As noted above the identification of the whole eligible cohort is currently under review, which is why non-core case are only included if they appear on any current data files.

### **Unknown outcome**

It will be noted that there were 69 pregnancies for whom we have been unable to date to obtain any information on delivery. This is in spite of repeated efforts using the birth notification system, searching through the gynaecological records for outcomes and writing to the mother herself. It is possible that some of these were social terminations of pregnancy, or that the mother left the area and has not been able to be identified since.

### **Version History**

Note that the version history was only started with the updates to version 4a.

#### Version 4a – March 2005

Variable MZ010, which fully identifies the size of the pregnancy, was added. Variables MZ013 & MZ014 were corrected so that they now tie up with KZ010 which was not previously the case and a similar variable for live births was added as MZ012. Variables MZ051 to MZ051k & MZ054, which related to Children in Focus attendance status, were removed because they were incorrect.

#### Version 5a – January 2009

- 5,247 eligible pregnancies not in the core sample of 14,541 were added. The total number of cases on the file is now therefore 19,788.
- For the additional non-core cases, variables MZ010 to MZ014 were set appropriately, variables MZ024a & MZ024b were filled in and variables MZ005a, MZ021a to MZ023a, MZ028a to MZ053 were set to -10.
- A variable indicating membership of the core ALSPAC sample of 14,541 pregnancies was introduced as MZ001.
- The file status variables MZ026a to MZ027d were removed as they are not guaranteed to be accurate and up to date.
- Missing categories 61 to 64 in variable MZ023 have been recoded to -8 to -5 so that missing values are all negative. This allows the previously merged -2 category to be split into -3 and -2 as per MZ023a.
- All bar a few basic descriptive variables have been set to -11 for triplet and quadruplet pregnancies.

## Variable Descriptions

### Membership of core sample

Variable MZ001 identifies the 14,541 pregnancies in the core ALSPAC sample.

#### mz001 In core ALSPAC sample

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	14541	73.5	73.5	73.5
2 No	5247	26.5	26.5	100.0
Total	19788	100.0	100.0	

### Mothers with more than one pregnancy in core sample

There were 657 pregnancies to women who had more than one pregnancy within the core ALSPAC sample. It should be noted that many of these pregnancies involved miscarriages and/or terminations and do not necessarily relate to more than one live born outcome of pregnancy. Note also that this variable is restricted to the core sample. There may well be women who had an eligible pregnancy not in the core sample as well as one or more pregnancies in the core sample.

#### mz005a More than 1 pregnancy in core ALSPAC sample

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 yes	657	3.3	4.5	4.5
2 no	13884	70.2	95.5	100.0
Total	14541	73.5	100.0	
Missing -10 Not in core sample	5247	26.5		
Total	19788	100.0		

It is possible to identify the pregnancies to women who had multiple pregnancies within the sample; 311 had two deliveries (3 of these also had a miscarriage), 332 had a delivery and a miscarriage (or termination) and 14 had two miscarriages. There is also a group of variables which will allow selection of only one pregnancy per mother, for example for genetic analyses, but access to these is only allowed with Jean Golding's permission.

## MZ File – Variable Descriptions

### Pregnancy size

In variable MZ010A, the outcome of pregnancy distinguishes between singleton and multiple pregnancies, with twins, triplets and quadruplets distinguished between in MZ010. No distinction is made between those multiple pregnancies for whom there was one survivor as opposed to those where there were no survivors or many survivors.

#### mz010 Pregnancy size

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Singleton	19462	98.4	98.7	98.7
2 Twin	252	1.3	1.3	100.0
3 Triplet	4	.0	.0	100.0
4 Quadruplet	1	.0	.0	100.0
Total	19719	99.7	100.0	
Missing -2 No delivery details	69	.3		
Total	19788	100.0		

#### mz010a Pregnancy size (summary)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Singleton	19462	98.4	98.7	98.7
2 Multiple	257	1.3	1.3	100.0
Total	19719	99.7	100.0	
Missing -2 No delivery details	69	.3		
Total	19788	100.0		

**Pregnancy outcome pre 20 weeks**

In MZ011, the detailed outcomes of pregnancy before 20 weeks are outlined. Miscarriage is defined as the death of a fetus in utero prior to delivery. Termination is deliberate delivery of a fetus, for whatever reason. However, fetuses that died in utero but for whom labour had to be induced are coded as miscarriages.

It can be seen that the miscarriages have been classified according to the gestation at which the miscarriage took place. For 18 pregnancies, there was no reliable information that allowed us to estimate the gestation at miscarriage and these are coded as miscarriage gestation not known.

Terminations were classified according to whether they were for a congenital malformation (code 6) or other reasons (code 7). Note that women intending to have a termination for social reasons were unlikely to take part in the study.

There was a miscellaneous group of 9 pregnancies with outcomes before 20 weeks that did not conform to the definition of either a miscarriage or termination - and these are coded separately (code 8). They include ectopic pregnancies and hydatidiform moles.

**mz011 Pregnancy outcome pre 20 wk**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 misc <12wks	246	1.2	1.2	1.2
	2 misc 12-15wks	160	.8	.8	2.1
	3 misc 16-19wks	51	.3	.3	2.3
	4 misc 20wks+	21	.1	.1	2.4
	5 misc gest N/K	18	.1	.1	2.5
	6 term foetal abnorm	47	.2	.2	2.8
	7 termin other reasons	20	.1	.1	2.9
	8 other	9	.0	.0	2.9
	10 delivered 20 wk+	19147	96.8	97.1	100.0
	Total	19719	99.7	100.0	
Missing	-2 No delivery details	69	.3		
Total		19788	100.0		

From MZ011, a summary variable relating to miscarriage in the first twenty weeks was identified by summing the first three codes. Thus the miscarriage which occurred definitely prior to 20 weeks is identified - but it may well be that the analyst wishes to combine this with those miscarriages for whom there is no recorded gestation (code 5 in MZ011). In which case they should carry out the appropriate recoding.

**mz011a Miscarriage < 20 wks**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes	457	2.3	2.3	2.3
	2 no	19262	97.3	97.7	100.0
	Total	19719	99.7	100.0	
Missing	-2 No delivery details	69	.3		
Total		19788	100.0		

## MZ File – Variable Descriptions

A further summary variable classifies termination of pregnancies for fetal problems (as opposed to maternal problems). This distinguishes between termination for congenital malformations and terminations for premature rupture of membranes.

### **mz011b Termination for fetal problem**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 yes, CM	49	.2	.2	.2
	3 yes, PRM	3	.0	.0	.3
	7 No	19667	99.4	99.7	100.0
	Total	19719	99.7	100.0	
Missing	-2 Outcome NK	69	.3		
Total		19788	100.0		

## **Survival**

In order to assist the appropriate sampling, three variables have been created that indicate whether the mother was eligible at various points to continue with the survey. MZ013 indicates survival of the child from 28 days onwards, and distinguishes between outcomes of a multiple pregnancy not all of whom survived (code 2), and outcomes of the pregnancy where all the children survived - this includes multiple and singleton pregnancy outcomes. In MZ014, the same variables are denoted for the one year survival rate and in MZ012 for survival at birth.

### **mz012 Live births**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No survivors	604	3.1	3.1	3.1
	2 1 survivor	1	.0	.0	3.1
	3 All survived	19114	96.6	96.9	100.0
	Total	19719	99.7	100.0	
Missing	-2 Outcome NK	69	.3		
Total		19788	100.0		

### **mz013 28-day survivors**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No survivors	649	3.3	3.3	3.3
	2 1 survivor	5	.0	.0	3.3
	3 All survived	19065	96.3	96.7	100.0
	Total	19719	99.7	100.0	
Missing	-2 Outcome NK	69	.3		
Total		19788	100.0		

### **mz014 1-year survivors**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No survivors	670	3.4	3.4	3.4
	2 1 survivor	7	.0	.0	3.4
	3 All survived	19042	96.2	96.6	100.0
	Total	19719	99.7	100.0	
Missing	-2 Outcome NK	69	.3		
Total		19788	100.0		



## MZ File – Variable Descriptions

### Season of conception

As a proxy for the season in which the child was conceived, MZ021A denotes the month of the last menstrual period. This was taken from the mother's report on enrolment, or where that was missing or impossible, computed from the date of delivery with adjustment made for the gestation at delivery.

It will be noted that April, May and June had lower numbers, this is due to the fact that the criteria for enrolment were conceptions which had an expected date of delivery between April 1991 and December 1992.

The year of the last menstrual period is given in MZ021B.

#### mz021a Month of last menstrual period

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1427	7.2	9.8	9.8
	2	1250	6.3	8.6	18.4
	3	1219	6.2	8.4	26.8
	4	699	3.5	4.8	31.6
	5	749	3.8	5.2	36.8
	6	819	4.1	5.6	42.4
	7	1245	6.3	8.6	51.0
	8	1296	6.5	8.9	59.9
	9	1405	7.1	9.7	69.6
	10	1514	7.7	10.4	80.0
	11	1383	7.0	9.5	89.5
	12	1526	7.7	10.5	100.0
	Total	14532	73.4	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-1 LMP not known	5	.0		
	Total	5256	26.6		
Total		19788	100.0		

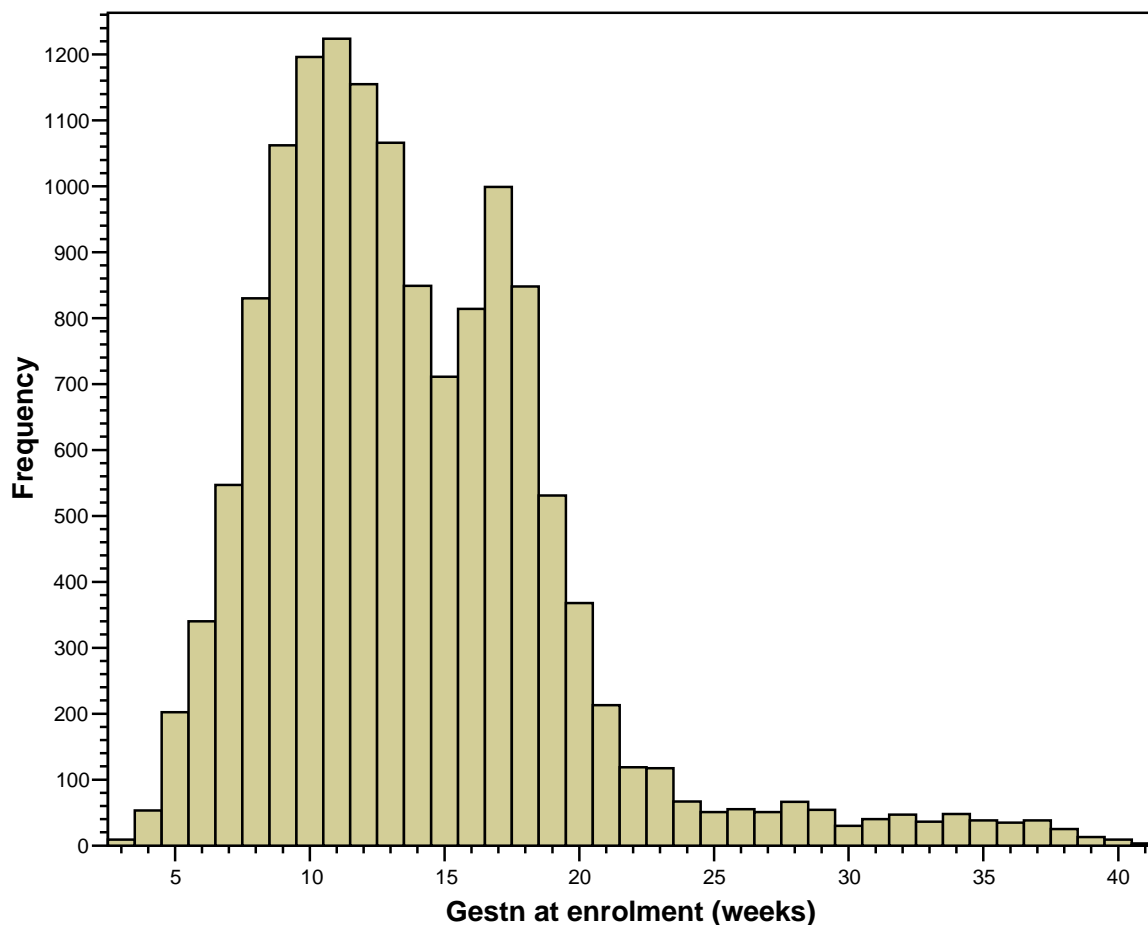
#### mz021b Year of last menstrual period

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1990	4008	20.3	27.6	27.6
	1991	8754	44.2	60.2	87.8
	1992	1770	8.9	12.2	100.0
	Total	14532	73.4	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-1 LMP not known	5	.0		
	Total	5256	26.6		
Total		19788	100.0		

**Gestation at enrolment**

In MZ022A, the gestation in completed weeks of the mother at enrolment is given. This has undergone a certain amount of cleaning to get rid of impossible dates. These included mothers who enrolled for one pregnancy that miscarried immediately and then became pregnant straight away and consequently appeared to have a negative enrolment period, keying errors in dates and other anomalies.

MZ022A relates only to the mothers who enrolled during the study pregnancy. There were 430 who enrolled after delivery and these are coded as -3. There were 8 mothers who enrolled apparently before conception and other categories include mothers who enrolled after they had miscarried (but who then completed a questionnaire), those where the enrolment date was not recorded and those where the LMP was not known.

**mz022a Gestn at enrolment (weeks)**

plus the following missing values:

**mz022a Gestn at enrolment (weeks)**

	Frequency	Percent
Missing -11 Triplet / quadruplet	5	.1
-10 Not in core sample	5246	90.0
-5 Enroll date NK	72	1.2
-4 Enrolled after misc	61	1.0
-3 Enrolled after delvy	432	7.4
-2 Enr before conceptn	8	.1
-1 LMP not known	5	.1
Total	5829	100.0

## MZ File – Variable Descriptions

For mothers who enrolled after the delivery of the child, the age of the baby at that enrolment was recorded in weeks. The majority (301) of these women were enrolled in the days after delivery - this was when the ALSPAC measurers were visiting the two large maternity hospitals and measuring the babies. Any mothers who were not members of ALSPAC already and who were there and willing at the time to enrol were enrolled in the study. Enrolment continued during the first year whenever mothers elicited interest in the study - but subsequently the computing difficulties in including late enrolment precluded any others being enrolled.

**mz022b Age of baby at enrolment (weeks)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	227	1.1	52.5	52.5
	1	74	.4	17.1	69.7
	2	28	.1	6.5	76.2
	3	21	.1	4.9	81.0
	4	15	.1	3.5	84.5
	5	13	.1	3.0	87.5
	6	16	.1	3.7	91.2
	7	18	.1	4.2	95.4
	8	6	.0	1.4	96.8
	9	3	.0	.7	97.5
	10	2	.0	.5	97.9
	15	1	.0	.2	98.1
	17	1	.0	.2	98.4
	18	1	.0	.2	98.6
	72	1	.0	.2	98.8
	92	1	.0	.2	99.1
	99	1	.0	.2	99.3
	150	1	.0	.2	99.5
	151	1	.0	.2	99.8
	165	1	.0	.2	100.0
	Total	432	2.2	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-2 Enrld in pregnancy	14105	71.3		
	Total	19356	97.8		
Total		19788	100.0		

**Gestation at delivery**

The gestational age that appears on MZ023 is the gestation when the pregnancy ended. In regard to miscarriages this will be the gestation at the time at which the fetus was thought to have died rather than the time at which the fetus was expelled. For all infants that were delivered after twenty weeks, if there was any suggestion that the gestation was preterm (i.e. less than 37 weeks) then the clinical records were reviewed by Professor Ian MacGillivray in the likelihood of the recorded gestation being correct or otherwise. Clinical records were reviewed not only for pregnancies that appeared preterm using the mother's LMP but also if the clinically recorded gestation based on ultrasound or the clinician's impression was less than 37 weeks or if the paediatric estimate of gestation was less than 37 weeks.

Details of whether or not the gestation was checked from clinical records are given in MZ022C. As previously indicated, cases which may have been preterm and deaths were checked in this way. In addition some attempt was made to assess all pregnancies that were said to be post term (i.e. 43 weeks or more). Again Professor Ian MacGillivray was responsible for this assessment.

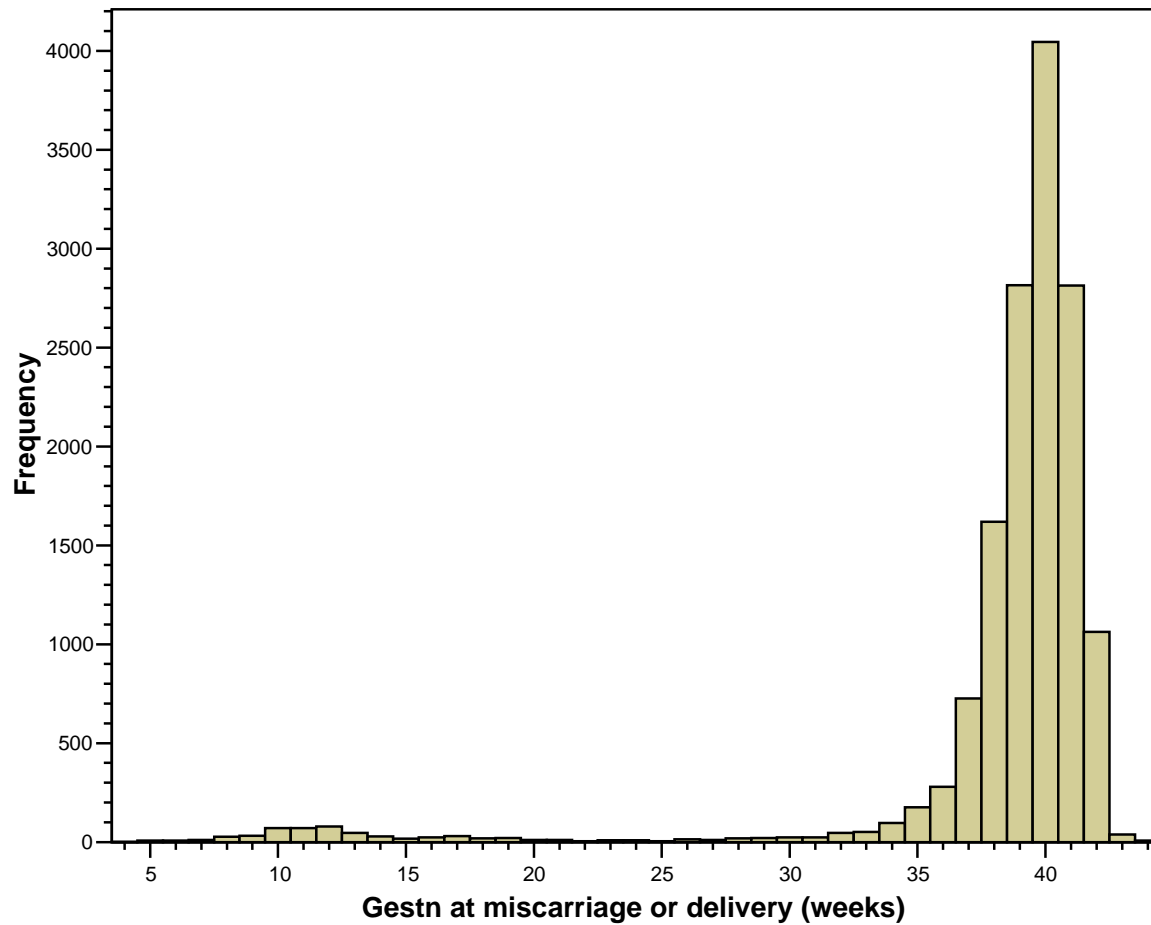
**mz022c Gestation checked with records**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 unchecked	12228	61.8	84.1	84.1
	1 checked	2309	11.7	15.9	100.0
	Total	14537	73.5	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	Total	5251	26.5		
Total		19788	100.0		

For 28 cases that delivered before 20 weeks gestation, the exact number of weeks of pregnancy were not recorded accurately, but there was a record which indicated which period of time the delivery had occurred within - these are coded as -8, -7, -6 and -5 in MZ023.

# MZ File – Variable Descriptions

## mz023 Gestn at miscarriage or delivery (weeks)



plus the following missing values:

## mz023 Gestn at miscarriage or delivery (weeks)

	Frequency	Percent
Missing -11 Triplet / quadruplet	5	.1
-10 Not in core sample	5246	97.8
-8 4-11	16	.3
-7 12-15	6	.1
-6 16-19	6	.1
-5 20-23	4	.1
-3 NK	12	.2
-2 No deliv dets	69	1.3
Total	5364	100.0

## MZ File – Variable Descriptions

The data in MZ023 were grouped in MZ023A for further ease of analysis.

### **mz023a grouped gestation**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 1-11	243	1.2	1.7	1.7
	2 12-15	176	.9	1.2	2.9
	3 16-19	98	.5	.7	3.6
	4 20-23	39	.2	.3	3.8
	5 24-27	37	.2	.3	4.1
	6 28-31	87	.4	.6	4.7
	7 32-36	650	3.3	4.5	9.2
	8 37-41	12018	60.7	83.1	92.3
	9 42-45	1108	5.6	7.7	100.0
	Total	14456	73.1	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-3 NK	12	.1		
	-2 no deliv dets	69	.3		
	Total	5332	26.9		
Total		19788	100.0		

## MZ File – Variable Descriptions

### Season of birth

In MZ024A the month of delivery of the pregnancy was recorded - regardless of outcome. In addition, in MZ024B the actual year of delivery is recorded.

#### mz024a Month of delivery

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1076	5.4	5.5	5.5
	2	869	4.4	4.4	9.9
	3	1044	5.3	5.3	15.2
	4	1829	9.2	9.3	24.5
	5	1948	9.8	9.9	34.4
	6	1883	9.5	9.6	44.0
	7	2002	10.1	10.2	54.1
	8	1888	9.5	9.6	63.7
	9	1827	9.2	9.3	73.0
	10	1819	9.2	9.2	82.3
	11	1762	8.9	9.0	91.2
	12	1728	8.7	8.8	100.0
	Total	19675	99.4	100.0	
Missing	-2 Miscarr, dates nk	44	.2		
	-1 Outcome NK	69	.3		
	Total	113	.6		
Total		19788	100.0		

#### mz024b Year of delivery

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1990	58	.3	.3	.3
	1991	8551	43.2	43.5	43.8
	1992	10964	55.4	55.7	99.5
	1993	102	.5	.5	100.0
	Total	19675	99.4	100.0	
Missing	-2 Miscarr, dates nk	44	.2		
	-1 Outcome NK	69	.3		
	Total	113	.6		
Total		19788	100.0		

## MZ File – Variable Descriptions

### Maternal age

The age of the mother was calculated from the mother's date of birth for specific time points. We therefore have variables MZ028A to MZ028E which relate to the age of the mother at, respectively, last menstrual period, delivery regardless of outcome of delivery, when the child would have been one year old, two years old and three years old.

In order to preserve confidentiality, the lower and upper age ranges have been grouped.

Note This is the age of the natural mother, and may not relate to the child's mother figure later in childhood.

**mz028a Grouped age of mother at LMP**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 15 < 16	43	.2	.3	.3
16	113	.6	.8	1.1
17	174	.9	1.2	2.3
18	286	1.4	2.0	4.2
19	336	1.7	2.3	6.6
20	417	2.1	2.9	9.4
21	493	2.5	3.4	12.8
22	624	3.2	4.3	17.1
23	782	4.0	5.4	22.5
24	937	4.7	6.4	28.9
25	1108	5.6	7.6	36.6
26	1135	5.7	7.8	44.4
27	1203	6.1	8.3	52.6
28	1185	6.0	8.2	60.8
29	1059	5.4	7.3	68.1
30	943	4.8	6.5	74.6
31	852	4.3	5.9	80.4
32	698	3.5	4.8	85.2
33	530	2.7	3.6	88.9
34	475	2.4	3.3	92.2
35	334	1.7	2.3	94.5
36	271	1.4	1.9	96.3
37	178	.9	1.2	97.6
38	133	.7	.9	98.5
39	89	.4	.6	99.1
40	73	.4	.5	99.6
41	27	.1	.2	99.8
42	15	.1	.1	99.9
43	11	.1	.1	99.9
44 >43	8	.0	.1	100.0
Total	14532	73.4	100.0	
Missing -11 Triplet / quadruplet	5	.0		
-10 Not in core sample	5246	26.5		
-3 LMP not known	5	.0		
Total	5256	26.6		
Total	19788	100.0		



# MZ File – Variable Descriptions

## mz028b Grouped age of mother at delivery

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	15 < 16	10	.1	.1	.1
	16	54	.3	.4	.5
	17	122	.6	.9	1.3
	18	184	.9	1.3	2.7
	19	285	1.4	2.1	4.7
	20	352	1.8	2.5	7.2
	21	406	2.1	2.9	10.2
	22	486	2.5	3.5	13.7
	23	639	3.2	4.6	18.3
	24	799	4.0	5.7	24.0
	25	949	4.8	6.8	30.8
	26	1084	5.5	7.8	38.6
	27	1101	5.6	7.9	46.6
	28	1152	5.8	8.3	54.9
	29	1083	5.5	7.8	62.7
	30	1031	5.2	7.4	70.1
	31	883	4.5	6.4	76.4
	32	774	3.9	5.6	82.0
	33	625	3.2	4.5	86.5
	34	495	2.5	3.6	90.1
	35	417	2.1	3.0	93.1
	36	290	1.5	2.1	95.1
	37	241	1.2	1.7	96.9
	38	156	.8	1.1	98.0
	39	109	.6	.8	98.8
	40	73	.4	.5	99.3
	41	54	.3	.4	99.7
	42	20	.1	.1	99.8
	43	9	.0	.1	99.9
	44 >43	13	.1	.1	100.0
	Total	13896	70.2	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-4 Outcome NK	69	.3		
	-2 Miscarried	572	2.9		
	Total	5892	29.8		
Total		19788	100.0		

# MZ File – Variable Descriptions

**mz028c Grpd age of mother 1 year after delvry**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	16 < 17	10	.1	.1	.1
	17	54	.3	.4	.5
	18	122	.6	.9	1.3
	19	184	.9	1.3	2.7
	20	285	1.4	2.1	4.7
	21	352	1.8	2.5	7.2
	22	406	2.1	2.9	10.2
	23	486	2.5	3.5	13.7
	24	639	3.2	4.6	18.3
	25	799	4.0	5.7	24.0
	26	949	4.8	6.8	30.8
	27	1084	5.5	7.8	38.6
	28	1101	5.6	7.9	46.6
	29	1152	5.8	8.3	54.9
	30	1083	5.5	7.8	62.7
	31	1031	5.2	7.4	70.1
	32	883	4.5	6.4	76.4
	33	774	3.9	5.6	82.0
	34	625	3.2	4.5	86.5
	35	495	2.5	3.6	90.1
	36	417	2.1	3.0	93.1
	37	290	1.5	2.1	95.1
	38	241	1.2	1.7	96.9
	39	156	.8	1.1	98.0
	40	109	.6	.8	98.8
	41	73	.4	.5	99.3
	42	54	.3	.4	99.7
	43	20	.1	.1	99.8
	44	9	.0	.1	99.9
	45 >44	13	.1	.1	100.0
	Total	13896	70.2	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not in core sample	5246	26.5		
	-4 Outcome NK	69	.3		
	-2 Miscarried	572	2.9		
	Total	5892	29.8		
Total		19788	100.0		

# MZ File – Variable Descriptions

## **mz028d Grpd age of mother 2 years after delvry**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 17 < 18	10	.1	.1	.1
18	54	.3	.4	.5
19	122	.6	.9	1.3
20	184	.9	1.3	2.7
21	285	1.4	2.1	4.7
22	352	1.8	2.5	7.2
23	406	2.1	2.9	10.2
24	486	2.5	3.5	13.7
25	639	3.2	4.6	18.3
26	799	4.0	5.7	24.0
27	949	4.8	6.8	30.8
28	1084	5.5	7.8	38.6
29	1101	5.6	7.9	46.6
30	1152	5.8	8.3	54.9
31	1083	5.5	7.8	62.7
32	1031	5.2	7.4	70.1
33	883	4.5	6.4	76.4
34	774	3.9	5.6	82.0
35	625	3.2	4.5	86.5
36	495	2.5	3.6	90.1
37	417	2.1	3.0	93.1
38	290	1.5	2.1	95.1
39	241	1.2	1.7	96.9
40	156	.8	1.1	98.0
41	109	.6	.8	98.8
42	73	.4	.5	99.3
43	54	.3	.4	99.7
44	20	.1	.1	99.8
45	9	.0	.1	99.9
46 >45	13	.1	.1	100.0
Total	13896	70.2	100.0	
Missing -11 Triplet / quadruplet	5	.0		
-10 Not in core sample	5246	26.5		
-4 Outcome NK	69	.3		
-2 Miscarried	572	2.9		
Total	5892	29.8		
Total	19788	100.0		

# MZ File – Variable Descriptions

## **mz028e Grpd age of mother 3 years after delvry**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 18 < 19	10	.1	.1	.1
19	54	.3	.4	.5
20	122	.6	.9	1.3
21	184	.9	1.3	2.7
22	285	1.4	2.1	4.7
23	352	1.8	2.5	7.2
24	406	2.1	2.9	10.2
25	486	2.5	3.5	13.7
26	639	3.2	4.6	18.3
27	799	4.0	5.7	24.0
28	949	4.8	6.8	30.8
29	1084	5.5	7.8	38.6
30	1101	5.6	7.9	46.6
31	1152	5.8	8.3	54.9
32	1083	5.5	7.8	62.7
33	1031	5.2	7.4	70.1
34	883	4.5	6.4	76.4
35	774	3.9	5.6	82.0
36	625	3.2	4.5	86.5
37	495	2.5	3.6	90.1
38	417	2.1	3.0	93.1
39	290	1.5	2.1	95.1
40	241	1.2	1.7	96.9
41	156	.8	1.1	98.0
42	109	.6	.8	98.8
43	73	.4	.5	99.3
44	54	.3	.4	99.7
45	20	.1	.1	99.8
46	9	.0	.1	99.9
47 >46	13	.1	.1	100.0
Total	13896	70.2	100.0	
Missing -11 Triplet / quadruplet	5	.0		
-10 Not in core sample	5246	26.5		
-4 Outcome NK	69	.3		
-2 Miscarried	572	2.9		
Total	5892	29.8		
Total	19788	100.0		

**Sub-studies**

There have been a number of sub-studies within and without ALSPAC, and these are outlined below.

**APIP**

This study which was run by Neil Marlow followed up children who had been delivered at less than 32 weeks gestation together with a number of matched controls. Parents within the APIP study were sent fewer questionnaires, on request from Neil Marlow, so that they would not be over-burdened. These included questionnaires at eight weeks and eight months.

**mz050 In APIP study**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	148	.7	1.0	1.0
2 No	14393	72.7	99.0	100.0
Total	14541	73.5	100.0	
Missing -10 Not in core sample	5247	26.5		
Total	19788	100.0		

**Aspirin randomised-control trial in pregnancy**

The CLASP study was a randomised multicentre controlled trial of low dose aspirin in women considered to be at high risk of developing pre-eclampsia and/or growth retardation.

MZ052 identifies women who were taking part in this study.

**mz052 In Clasp study**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	42	.2	.3	.3
2 No	14499	73.3	99.7	100.0
Total	14541	73.5	100.0	
Missing -10 Not in core sample	5247	26.5		
Total	19788	100.0		

**The thermometer study**

The thermometer study was designed to determine whether possession of a nursery thermometer was beneficial to the health of the child.

**mz053 In Thermometer study**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Yes	2653	13.4	18.2	18.2
2 No	11888	60.1	81.8	100.0
Total	14541	73.5	100.0	
Missing -10 Not in core sample	5247	26.5		
Total	19788	100.0		