### THE ALSPAC STUDY

## **PUB8 FILE**

### DATA COLLECTED FROM THE QUESTIONNAIRE

# Growing and Changing (8)

### At 192 Months

### **Prepared by**

### The ALSPAC Study Team

Documentation giving frequencies, background and instructions for use.

Last updated for version 1a of the built file.

February 2010

© University of Bristol

# **Contents**

Introduction			Page 3
Administrative var	riables		Page 5
Section A		Text Editing Notes & Frequencies	Page 13 Page 19
Sections B & C (F	emale)	Text Editing Notes & Frequencies	Page 33 Page 36
Sections B & C (N	/lale)	Text Editing Notes & Frequencies	Page 37 Page 40
Section D		Text Editing Notes & Frequencies	Page 41 Page 43
Appendix A	Gene	eral Coding Instructions	Page 47
Appendix B	Spec	ific Coding Instructions	Page 50
Appendix C	Ques	stionnaire Methodology	Page 55

### Introduction

#### Contents

The eighth in the series of puberty questionnaires was sent out to the study teenager when they were 192 months old. "Growing and Changing (7)" was tied to the TeenFocus 3 clinic and was adapted accordingly. "Growing and Changing (8)" reverted to being sent out on the usual rolling basis by age and is the same as "Growing and Changing (6)".

#### **Male Tanner Genital Staging**

As a consequence of an ongoing project conducted by Carol Rubin, serious problems with the male Tanner genital stage data have come to light. When the data from the first five puberty questionnaires were analyzed longitudinally, it was found that 27% of males went backwards in genital stage. This is in contrast with 3-4% going backwards for each of male pubic hair stage, female breast stage and female pubic hair stage. In addition, even after exclusion of males who go backwards in genital stage and males less than 10 years old, the estimated ages at transition into Tanner genital stages 2 or 3 produced by the modeling process are at least a year earlier than expected. It is strongly recommended that the male Tanner genital stage data (variable PUB850 on this file) are not used.

#### **Questionnaire versions**

There was only one version of each of the male and female copies of this questionnaire.

#### Sample & response rates

There are a total of 15,151 records on this built file. This number is made up of the 14,676 fetuses in the core ALSPAC sample (regardless of whether or not this questionnaire was sent out to them or whether it was returned) plus 475 eligible children not in the core sample to whom questionnaires were sent out. Note that questionnaires were completed by 207 of these 475 children.

Of the 14,676 fetuses in the core ALSPAC sample, 14,062 were live born. This questionnaire was sent out to 9,534 (68%) of these live born children. As of 31<sup>st</sup> January 2010 completed questionnaires had been returned by 4,553 (48%) of these children, which is 32% of the 14,062 live born children. For further information on the ALSPAC sample, please see section 5 of the "Guide to ALSPAC data" which can be found in the "Collaborator Pack" on the ALSPAC documentation CD.

Note that 5 of the 4,760 completed questionnaires belong to children from triplet or quadruplet pregnancies, all of whom are in the core sample. For reasons of confidentiality the data for these children are not available. The administrative variables PUB881-PUB887 remain, but PUB887a has been set to 2 "No" and all other variables have been set to –11.

#### Format of this documentation

The bulk of this documentation consists of the text of the questionnaire interspersed with editing notes and frequency tables of the variables on the data file.

### PUB8 File - Introduction

## **Built file version history**

Built version 1a – February 2010 The first version of the built file.

## **Administrative variables**

### pub881 Questionnaire sent

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	10009	66.1	66.1	66.1
	2 No	5142	33.9	33.9	100.0
	Total	15151	100.0	100.0	

#### pub882 Reminder sent

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	6598	43.5	65.9	65.9
	2 No	3411	22.5	34.1	100.0
	Total	10009	66.1	100.0	
Missing	-2 Questionnaire not sent	5142	33.9		
Total		15151	100.0		

### pub885 Questionnaire return status (as of 31/01/10)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Not returned	5213	34.4	52.1	52.1
	2 Returned, completed	4760	31.4	47.6	99.6
	3 Returned, blank	36	.2	.4	100.0
	Total	10009	66.1	100.0	
Missing	-2 Questionnaire not sent	5142	33.9		
Total		15151	100.0		

### pub886 Questionnaire returned (as of 31/01/10)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	4796	31.7	47.9	47.9
	2 No	5213	34.4	52.1	100.0
	Total	10009	66.1	100.0	
Missing	-2 Questionnaire not sent	5142	33.9		
Total		15151	100.0		

#### pub887 Questionnaire completed (as of 31/01/10)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	4760	31.4	47.6	47.6
	2 No	5249	34.6	52.4	100.0
	Total	10009	66.1	100.0	
Missing	-2 Questionnaire not sent	5142	33.9		
Total		15151	100.0		

#### pub887a Data available (as of 31/01/10)

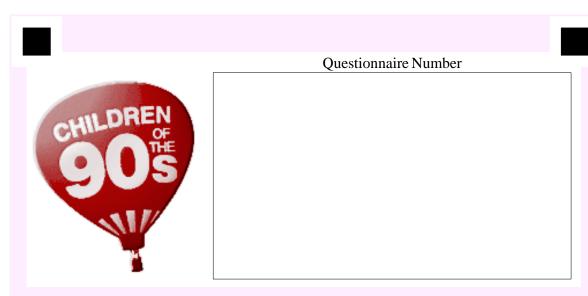
		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	4755	31.4	31.4	31.4
	2 No	10396	68.6	68.6	100.0
	Total	15151	100.0	100.0	

### PUB8 File - Introduction

### pub888 Questionnaire version

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Female v ersion 1 - 25/05/07	2814	18.6	59.2	59.2
	2 Male version 1 - 24/05/07	1941	12.8	40.8	100.0
	Total	4755	31.4	100.0	
Missing	-11 Triplet / quadruplet	5	.0		
	-10 Not completed	10391	68.6		
	Total	10396	68.6		
Total		15151	100.0		

#### PAGE 1 OF THE FEMALE QUESTIONNAIRE



# Female Teenager questionnaire

# **GROWING AND CHANGING (8)**

You may have seen this type of questionnaire before.

There are important changes to a teenager's body that can happen as early as 6 or as late as 20. At this time of life we have some questions we need to ask and which we need to hear about at regular intervals.

Use a **BLACK** or **BLUE** pen when you fill in the questionnaire.

Please read carefully the instructions on page 3.

25.05.07

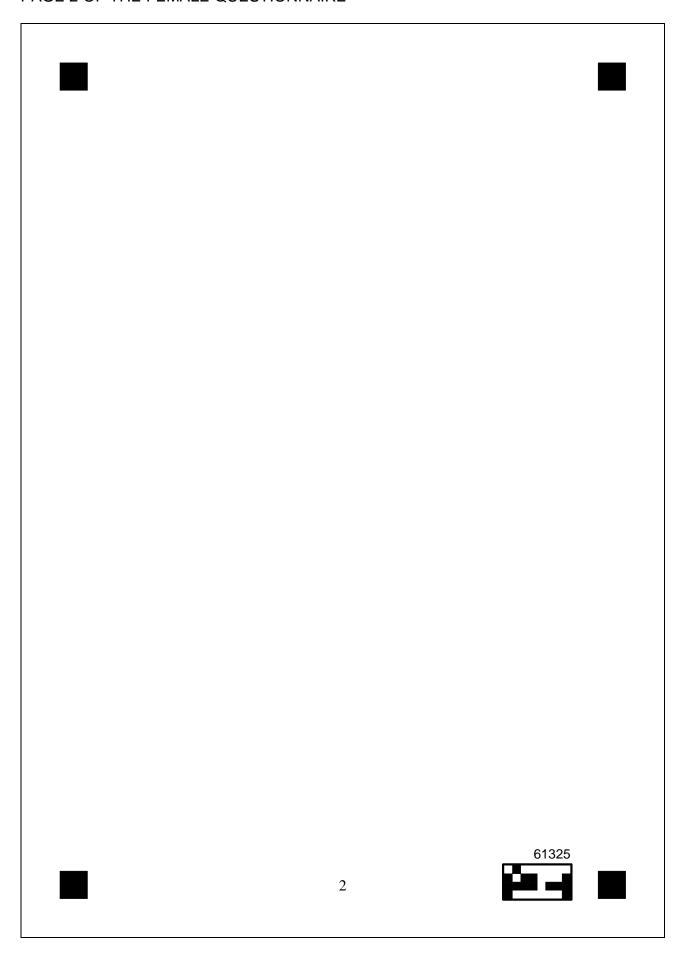




1

### PUB8 File - Introduction

### PAGE 2 OF THE FEMALE QUESTIONNAIRE



#### PAGE 3 OF THE FEMALE QUESTIONNAIRE

You will have seen fairly recently that we now have an electronic scanner that will record your answers automatically.

So, please mark your answers with a cross like this:

If you make a mistake, shade the box in like this:



If you are writing numbers make sure they are inside the box like this: 6

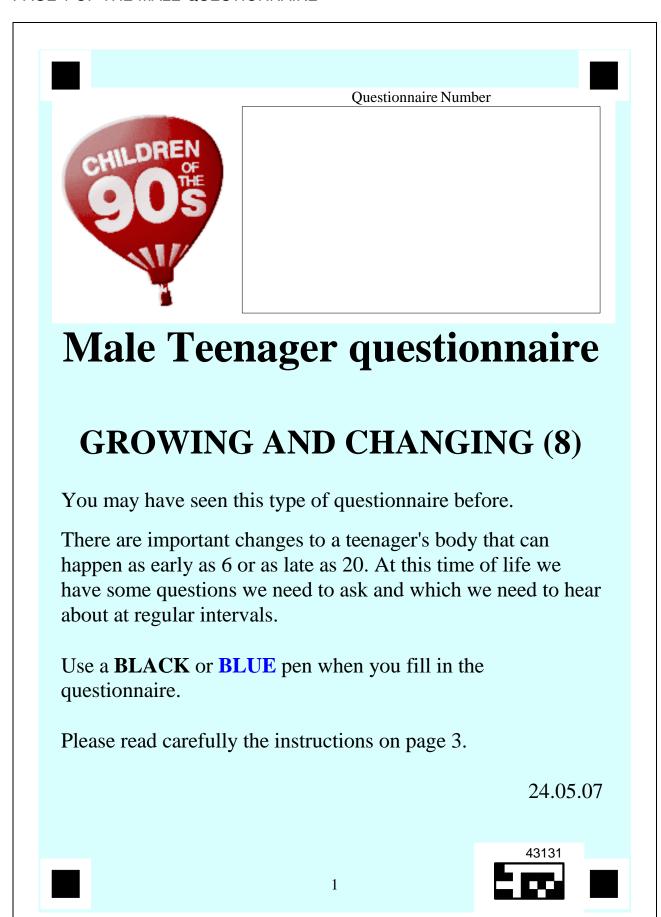
Use a **BLACK** or **BLUE** pen

Your answers will of course still be completely confidential and not attached to your actual name

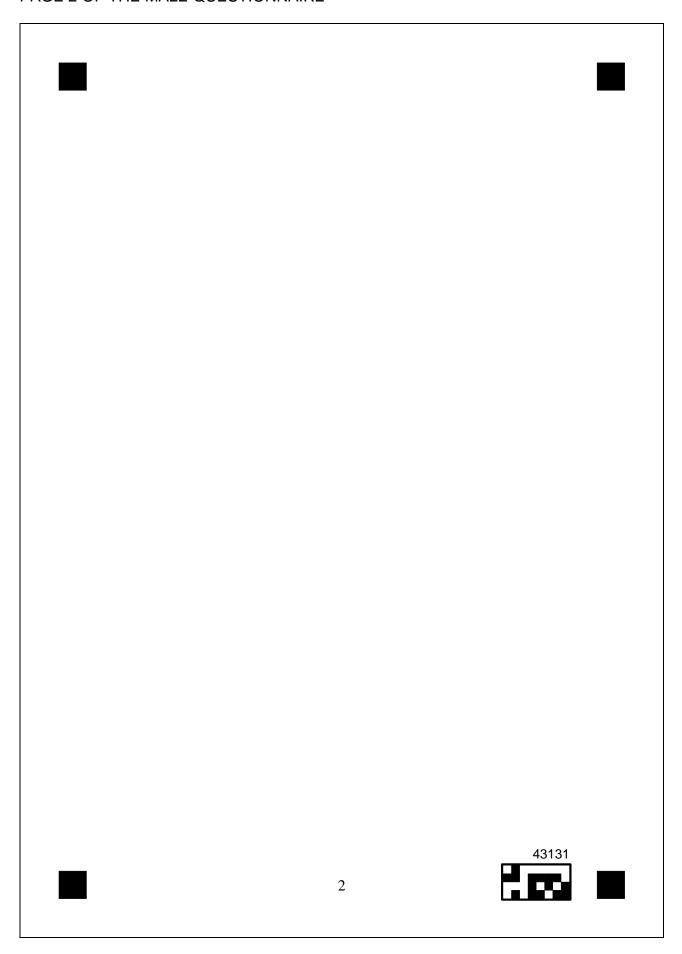


3

#### PAGE 1 OF THE MALE QUESTIONNAIRE



### PAGE 2 OF THE MALE QUESTIONNAIRE



#### PAGE 3 OF THE MALE QUESTIONNAIRE

You will have seen fairly recently that we now have an electronic scanner that will record your answers automatically.

So, please mark your answers with a cross like this:

If you make a mistake, shade the box in like this:



Then cross the correct box.

If you are writing numbers make sure they are inside the box like this:

6

Use a **BLACK** or **BLUE** pen

Your answers will of course still be completely confidential and not attached to your actual name

### PAGE 4 OF THE FEMALE QUESTIONNAIRE

wa	hat is your height (withou ne best way to measure he all, and then ask someone ad, and to measure the dis	<b>ight</b> is to stand b to make a mark	on the wall at the highest	_
	feet inche	s OR	metres cer	atimetres
	hat is your weight (withou			
	stones pound	S OR	kilos	
	the past month, what was			
_	robics)?	such as rummig,	dance, gymnastics, netoa	ii, swimming, or
_			4-6 times a week	11, swimming, or  4 □
_	robics)?	_	-	_
_	robics)?	1 🗆	4-6 times a week	4 🗆
_	none less than once a week	1	4-6 times a week	4 🗆

### PAGE 5 OF THE FEMALE QUESTIONNAIRE

	ve you started your p	-	► If no plo	aga ga ta	A 10 on nogo 7
	Yes ¹□ N	0 2 🔲 ——	<b>→</b> 11 <u>110,</u> pre	ase go to i	A10 on page 7
If <u>yes</u> ,					
a)	How old were you	when you had	l your first pe	riod?	
		years old			
A5. Wl	nen exactly was your	first period?			
	month	yea	r		
A6. a)	In the <b>past year</b> , he each period?	ow many <b>day</b>	s of bleeding	have you	usually had during
	Number of days		It varies	88 🔲	Don't know 99 □
b)	If you don't know,	or if it varie	es is it probab	ly:	
	3 days or less	1 🔲			
	4-6 days	2 🔲			
	7 days or more	3 🗖			
words,	the past year, how reference the past year, how many days were kt period?				en periods? In other iod to the <u>first day</u> of
	Number of days		It varied	88 🔲	Don't know 99 □

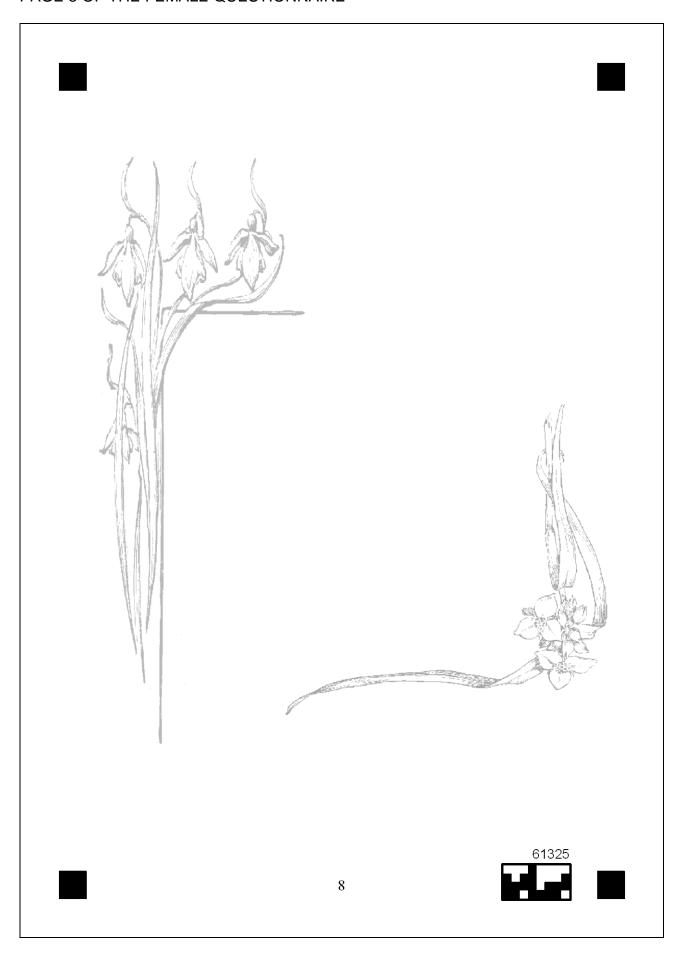
### PAGE 6 OF THE FEMALE QUESTIONNAIRE

<b>A</b> 8. Ha	ave you <b>ever</b> had any of the following <b>symptoms</b> associated with <b>your period</b> ?
a)	Heavy or prolonged <b>bleeding</b>
	Yes ¹□ No ²□ — If <u>no</u> , go to A8b below
f <u>yes</u> ,	i) did you contact a doctor for this?
	Yes 1 No 2
b)	Severe <b>cramps</b> with your period?
	Yes ¹□ No ²□
If <u>yes</u> ,	i) did you contact your doctor for this?
	Yes ¹□ No ²□
c)	Period-type pains or pain in your <b>pelvic</b> area (lower part of your tummy) for most days of the month even when you are not bleeding?
	Yes 1 No 2
If <u>yes</u> ,	
	i) Did you contact your doctor for this?
	Yes ¹□ No ²□
	61325

### PAGE 7 OF THE FEMALE QUESTIONNAIRE

		amps, their C	GP may pr	rescribe the	heir periods e.g. heavy b oral contraceptive pill ( gen pills') to help.	
	ave you st 12 mo		ontraceptiv	es or birth	control pills, for any rea	ason during the
	Yes	1 🗖	No	2 🔲	Don't know 9 □	
A10.		Has a doctor e hyroid medic			had a <b>thyroid problem</b>	or asked you to
	Yes	1 🔲	No	2 🔲	Don't know 9 □	
If	<u>yes</u> ,					
	b) V	Vhat kind of	thyroid pr	oblem did	the doctor say you had?	
	b) V	Vhat kind of	thyroid pi	roblem did	the doctor say you had?	
A11.					the doctor say you had?	
A11.			o have hai			
A11.	Have	you started to	o have hai	ir growing i		
A11.	Have	you started to	o have hai	ir growing i		
A11.	Have	you started to	o have hai	ir growing i		
A11.	Have	you started to	o have hai	ir growing i		
A11.	Have	you started to	o have hai	ir growing i		

### PAGE 8 OF THE FEMALE QUESTIONNAIRE



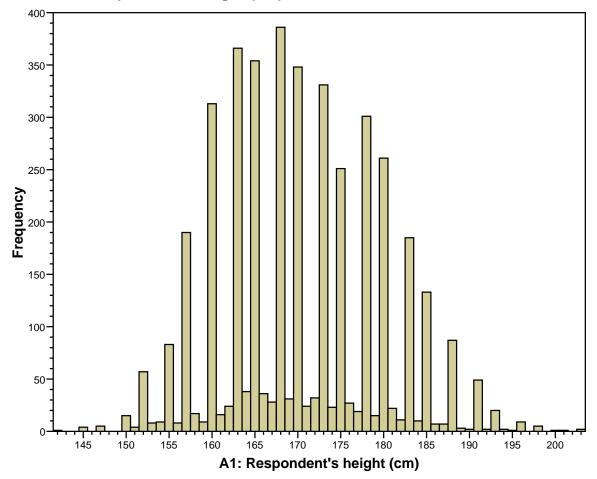
### PAGE 4 OF THE MALE QUESTIONNAIRE

Section A				
A1. What is your height (without The best way to measure he wall, and then ask someon head, and to measure the decrease.)	<b>neight</b> is to sta e to make a m	ark on th	ne wall at the highest	_
feet inche	s	OR	metres	centimetres
A2. What is your weight (with Please fill in using kilos or				
stones por	ands	OR	kilos	
A3. In the past month, what wa <b>vigorous</b> physical activity	_			•
none	1		4-6 times a week	4 🔲
less than once a week	2 🔲		daily	5 🔲
1-3 times a week	3			
A4. Has your voice changed at	all?			
no it is the same	1		yes, occasionally it is a lot lower	2 🔲
yes, it has now changed totally	3		not sure	4 🔲
A5. Have you started to have h	nair growing in	n the arm	pits?	
Yes ¹□	No 2			
			F	43131

[Editing: Twenty-three reported imperial heights for which the number of inches did not correspond to the number of feet (e.g. 5ft 79in) were set to -1. Four reported imperial measurements with values only in the inches box (in the valid range of 0-11) were set to -1, as were nineteen other outlying imperial measurements. For thirty-four imperial heights for which the inches box was left blank (sixteen of 5ft and eighteen of 6ft) the inches variable was recoded to 0. Nine reported metric measurements with values only in the cm box (in the valid range of 0-99) were set to -1, as were twenty-two other outlying metric measurements.

A single variable for height in cm was then created using a conversion factor of 1in = 2.54cm for imperial measurements (rounding to the nearest integer). If both imperial and metric heights were reported then the metric height was taken unless the two heights were not equivalent and the imperial height was closer to the median (5ft 10in for males and 5ft 5in for females).]

pub803 A1: Respondent's height (cm)



plus the following missing values:

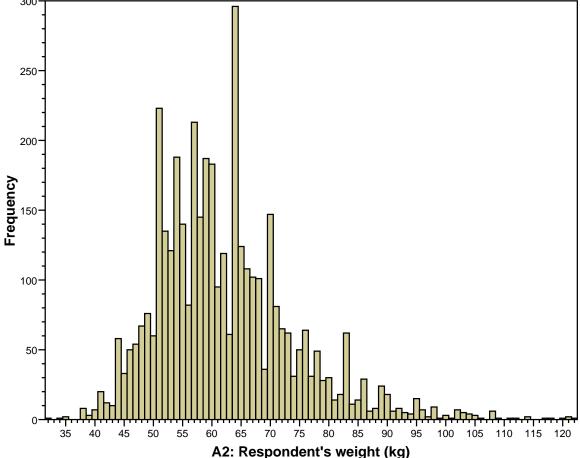
pub803 A1: Respondent's height (cm)

		Frequency	Percent
Missing	-1 No response	562	100.0

[Editing: Forty reported imperial weights for which the number of pounds did not correspond to the number of stones (e.g. 8st 95lb) were set to -1. Three reported imperial measurements with values only in the pounds box (in the valid range of 0 to 13) were set to -1, as was one reported imperial measurements over 28st. For two hundred and forty-eight imperial weights for which the pounds box was left blank the pounds variable was recoded to 0.

Three reported metric measurements less than 30kg were set to -1. A single variable for weight in kg was then created using a conversion factor of 1lb = 0.4536kg for imperial measurements (rounding to the nearest integer). If both imperial and metric weights were reported then the metric weight was taken unless the two weights were not equivalent and the imperial weight was closer to the median (65kg for males and 57kg for females).]

pub804 A2: Respondent's weight (kg)



plus the following missing values and outliers > 122:

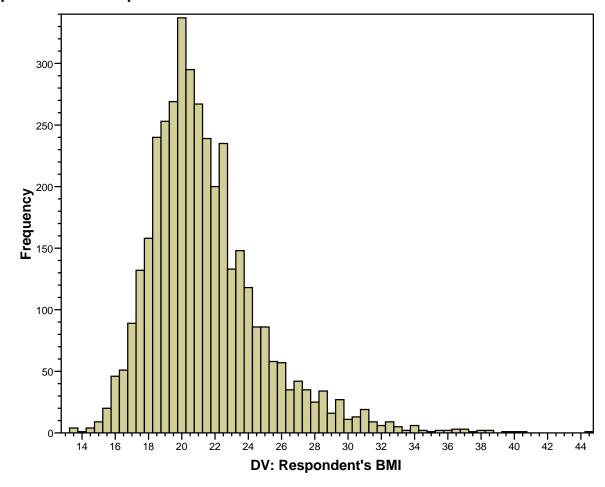
pub804 A2: Respondent's weight (kg)

partition partition and the state of the sta					
		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	133	2	.3	100.0	100.0
Missing	-1 No response	692	99.7		
Total		694	100.0		

### **Derived variable - BMI**

Body Mass Index (BMI) is calculated as PUB805 from PUB803 and PUB804 as weight (in kg) divided by height (in m) squared rounded to 1 decimal place. Missing values of -6 were copied across and if either component had value -1 then the BMI was also set to -1.

pub805 DV: Respondent's BMI



plus the following missing values:

pub805 DV: Respondent's BMI

		Frequency	Percent
Missing	-1.0 Missing	904	100.0

pub809 A3: Average number of times respondent participated in vigorou activity in past month

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 None	301	6.3	6.4	6.4
	2 <1 per week	790	16.6	16.8	23.2
	3 1-3 per week	2117	44.5	45.0	68.2
	4 4-6 per week	942	19.8	20.0	88.3
	5 Daily	552	11.6	11.7	100.0
	Total	4702	98.9	100.0	
Missing	-1 Not stated	53	1.1		
Total		4755	100.0		

### **Question A4 (male)**

### pub860 A4 (male): Respondent's voice has changed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No	252	5.3	13.2	13.2
	2 Yes, occasionally a lot lower	301	6.3	15.7	28.9
	3 Yes, has changed totally	1165	24.5	60.8	89.7
	4 Not sure	198	4.2	10.3	100.0
	Total	1916	40.3	100.0	
Missing	-7 Respondent is female	2814	59.2		
	-1 Not stated	25	.5		
	Total	2839	59.7		
Total		4755	100.0		

### **Question A4 (female)**

[Editing: Thirty-five respondents left the initial part of A4 blank but reported an age in A4a and went on to make a number of relevant responses to questions A5 to A9. The initial part of A4 was therefore recoded to 1 for all of these cases. Then if the variable for the initial part of A4 had value -1 the variables for A4 to A9 were all set to -6 and if the variable for the initial part of A4 had value 2 the variables for A4a to A9 were all set to -2.]

pub810 A4 (female): Respondent has started menstrual periods

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	2796	58.8	99.4	99.4
	2 No	16	.3	.6	100.0
	Total	2812	59.1	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	Total	1943	40.9		
Total		4755	100.0		

pub811 A4a: Age of respondent when first period occurred

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	8	2	.0	.1	.1
	9	9	.2	.3	.4
	10	111	2.3	4.1	4.5
	11	487	10.2	17.9	22.4
	12	735	15.5	27.1	49.5
	13	851	17.9	31.3	80.8
	14	416	8.7	15.3	96.1
	15	93	2.0	3.4	99.6
	16	12	.3	.4	100.0
	Total	2716	57.1	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	-1 Not stated	80	1.7		
	Total	2039	42.9		
Total		4755	100.0		

### **Question A5 (female)**

[Editing: A variable was then created as PUB813c indicating which components of the date were missing. Then if a year had been reported but not a month the month was imputed as 6 unless the reported year equalled the year of receipt in which case the month was set to the month of receipt divided by 2 and rounded to nearest integer. If a month had been reported but not a year then the month was recoded to -1.]

pub813a A5: Date of first period - month

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 January	216	4.5	9.3	9.3
	2 February	151	3.2	6.5	15.7
	3 March	132	2.8	5.7	21.4
	4 April	175	3.7	7.5	28.9
	5 May	179	3.8	7.7	36.6
	6 June	311	6.5	13.3	49.9
	7 July	184	3.9	7.9	57.8
	8 August	323	6.8	13.8	71.6
	9 September	188	4.0	8.1	79.7
	10 October	156	3.3	6.7	86.4
	11 Nov ember	166	3.5	7.1	93.5
	12 December	152	3.2	6.5	100.0
	Total	2333	49.1	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	-1 Not stated	463	9.7		
	Total	2422	50.9		
Total		4755	100.0		

pub813b A5: Date of first period - year

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	2000	5	.1	.2	.2
	2001	41	.9	1.8	2.0
	2002	200	4.2	8.6	10.5
	2003	450	9.5	19.3	29.8
	2004	613	12.9	26.3	56.1
	2005	566	11.9	24.3	80.4
	2006	345	7.3	14.8	95.2
	2007	92	1.9	3.9	99.1
	2008	20	.4	.9	100.0
	2009	1	.0	.0	100.0
	Total	2333	49.1	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	-1 Not stated	463	9.7		
	Total	2422	50.9		
Total		4755	100.0		

### PUB8 File - Section A

pub813c DV: Components missing (after identification of illegal dates)

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	0 None	2203	46.3	78.8	78.8
	1 Month only	130	2.7	4.6	83.4
	2 Year only	25	.5	.9	84.3
	3 Month and year	438	9.2	15.7	100.0
	Total	2796	58.8	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	Total	1959	41.2		
Total		4755	100.0		

#### Derived variable - Age at first period

The age of the respondent at her first period was calculated in months as PUB812 from the date of first period and the date of birth from ALSPAC's central database. For the purposes of this calculation the day of first period was assumed to be 15 unless the first period occurred in the same month & year as the questionnaire was completed, in which case it was set to the day of completion divided by 2 and rounded to nearest integer.

If the month of first period was the same as the month of birth then PUB812 was checked against the reported age in years from A4a (PUB811) and manual adjustment made to ensure compatibility if possible. For example, if PUB811 was 12 and the month of first period was the same as the month of birth then values of 143 in PUB812 were recoded to 144 and values of 156 recoded to 155. However, some inconsistencies remain between PUB811 and PUB812 because the reported values are not compatible.

110 90 80 70 40 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 DV: Age of respondent at first period (months)

PUB812 DV: Age of respondent at first period (months)

plus the following missing values:

pub812 DV: Age of respondent at first period (months)

•	•		•
		Frequency	Percent
Missing	-7 Respondent is male	1941	80.1
	-6 Questions A4 to A9 omitted	2	.1
	-2 Periods not started	16	.7
	-1 Missing	463	19.1
	Total	2422	100.0

[Editing: The information from part a of this question was captured as three separate variables with some instance of multiple entries. For example, one hundred and four respondents reported a number of days and also ticked "It varies". The information was combined into a single variable PUB815 by taking the reported number of days and recoding values of -1 to -3 if "It varies" had been ticked and then recoding values of -1 to -9 if "Don't know" had been ticked. So a reported number of days was taken in preference to either tick and "It varies" in preference to "Don't know".]

pub815 A6a: Usual number of days of bleeding during each of the respondent's per the past year

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	0	5	.1	.3	.3
	1	1	.0	.1	.3
	2	9	.2	.5	.8
	3	90	1.9	4.7	5.4
	4	318	6.7	16.5	21.9
	5	716	15.1	37.1	59.0
	6	384	8.1	19.9	78.9
	7	347	7.3	18.0	96.8
	8	37	.8	1.9	98.8
	9	11	.2	.6	99.3
	10	10	.2	.5	99.8
	14	3	.1	.2	100.0
	Total	1931	40.6	100.0	
Missing	-9 Don't know	106	2.2		
	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-3 It varies	727	15.3		
	-2 Periods not started	16	.3		
	-1 Not stated	32	.7		
	Total	2824	59.4		
Total		4755	100.0		

pub816 A6b: Estimated usual number of days of bleeding during each of the respor periods in the past year

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 3 or less	64	1.3	6.3	6.3
	2 4 to 6	826	17.4	81.9	88.3
	3 7 or more	118	2.5	11.7	100.0
	Total	1008	21.2	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	-1 Not stated	1788	37.6		
	Total	3747	78.8		
Total		4755	100.0		

[Editing: The information from this question was captured as three separate variables with some instance of multiple entries. For example, forty-five respondents reported a number of days and also ticked "It varies". The information was combined into a single variable PUB817 by taking the reported number of days and recoding values of -1 to -3 if "It varies" had been ticked and then recoding values of -1 to -9 if "Don't know" had been ticked. So a reported number of days was taken in preference to either tick and "It varies" in preference to "Don't know".]

A7: Usual length of respondent's menstrual cycle in the past year

pub817 A7: Usual length of respondent's menstrual cycle in the past year

plus the following missing values and outliers > 63:

pub817 A7: Usual length of respondent's menstrual cycle in the past year

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	90	1	.0	50.0	50.0
	120	1	.0	50.0	100.0
	Total	2	.1	100.0	
Missing	-9 Don't know	407	11.5		
	-7 Respondent is male	1941	54.6		
	-6 Questions A4 to A9 omitted	2	.1		
	-3 It varies	1125	31.7		
	-2 Periods not started	16	.5		
	-1 Not stated	61	1.7		
	Total	3552	99.9		
Total		3554	100.0		

[Editing: For each of parts a, b and c if the initial indicator variable had value 2 then the variable for (i) was set to -3 and if the initial indicator variable had value -1 then both variables for that part were set to -5.]

pub820 A8a: Respondent has had heavy or prolonged bleeding associated with per

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	1007	21.2	36.5	36.5
	2 No	1751	36.8	63.5	100.0
	Total	2758	58.0	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8a omitted	38	.8		
	-2 Periods not started	16	.3		
	Total	1997	42.0		
Total		4755	100.0		

pub821 A8a1: Doctor contacted about period associated heavy or prolonged bleedir

•		•		, ,	J
		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	268	5.6	26.6	26.6
	2 No	739	15.5	73.4	100.0
	Total	1007	21.2	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8a omitted	31	.7		
	-3 No heavy bleeding	1747	36.7		
	-2 Periods not started	16	.3		
	-1 Not stated	11	.2		
	Total	3748	78.8		
Total		4755	100.0		

### PUB8 File - Section A

pub822 A8b: Respondent has had severe cramps associated with period

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	1516	31.9	56.3	56.3
	2 No	1175	24.7	43.7	100.0
	Total	2691	56.6	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8b omitted	105	2.2		
	-2 Periods not started	16	.3		
	Total	2064	43.4		
Total		4755	100.0		

pub823 A8b1: Doctor contacted about period associated severe cramps

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	328	6.9	21.7	21.7
	2 No	1186	24.9	78.3	100.0
	Total	1514	31.8	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8b omitted	98	2.1		
	-3 No severe cramps	1174	24.7		
	-2 Periods not started	16	.3		
	-1 Not stated	10	.2		
	Total	3241	68.2		
Total		4755	100.0		

### PUB8 File - Section A

pub824 A8c: Respondent has had period-type pain(s) in pelvic area most days of meven when not bleeding

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	547	11.5	20.7	20.7
	2 No	2096	44.1	79.3	100.0
	Total	2643	55.6	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8c omitted	153	3.2		
	-2 Periods not started	16	.3		
	Total	2112	44.4		
Total		4755	100.0		

### pub825 A8c1: Doctor contacted about period-type pain in pelvic area occurring when not blee

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	127	2.7	23.0	23.0
	2 No	425	8.9	77.0	100.0
	Total	552	11.6	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-5 Question A8c omitted	140	2.9		
	-3 No period-type pains when not bleeding	2094	44.0		
	-2 Periods not started	16	.3		
	-1 Not stated	10	.2		
	Total	4203	88.4		
Total		4755	100.0		

pub827 A9: Respondent has taken oral contraceptives/birth control pills in past yea

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	695	14.6	25.2	25.2
	2 No	2045	43.0	74.1	99.3
	9 Don't know	19	.4	.7	100.0
	Total	2759	58.0	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-6 Questions A4 to A9 omitted	2	.0		
	-2 Periods not started	16	.3		
	-1 Not stated	37	.8		
	Total	1996	42.0		
Total		4755	100.0		

### **Question A10**

# pub828 A10a: Doctor has advised respondent of Thyroid problem or to take Thyroid medicine/treatment

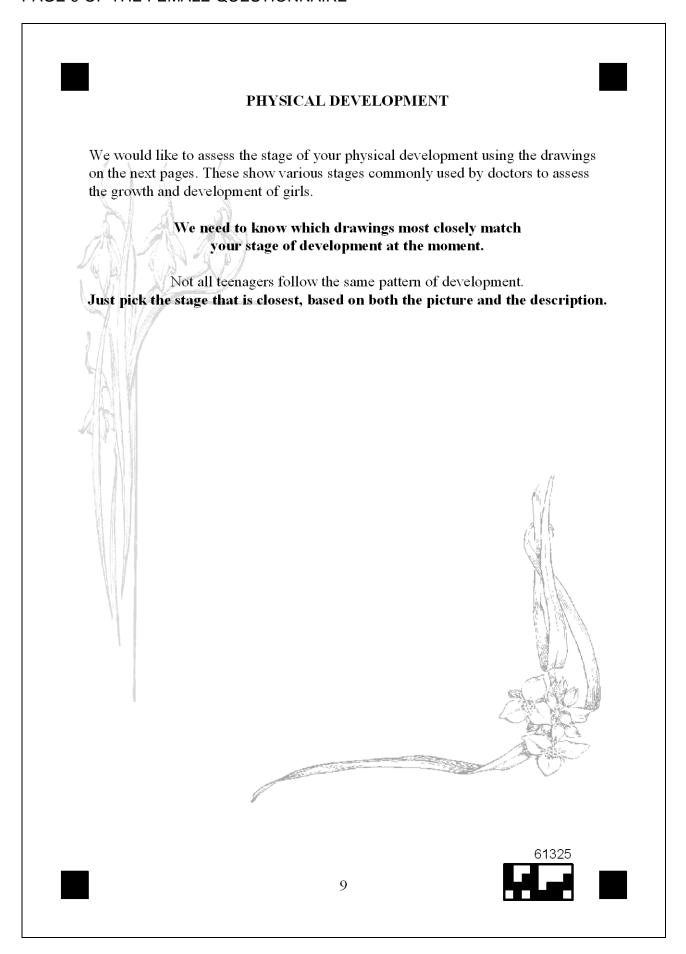
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	20	.4	.7	.7
	2 No	2688	56.5	97.7	98.5
	9 Don't know	42	.9	1.5	100.0
	Total	2750	57.8	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-1 Not stated	64	1.3		
	Total	2005	42.2		
Total		4755	100.0		

### Question A11 (female) / A4 (male)

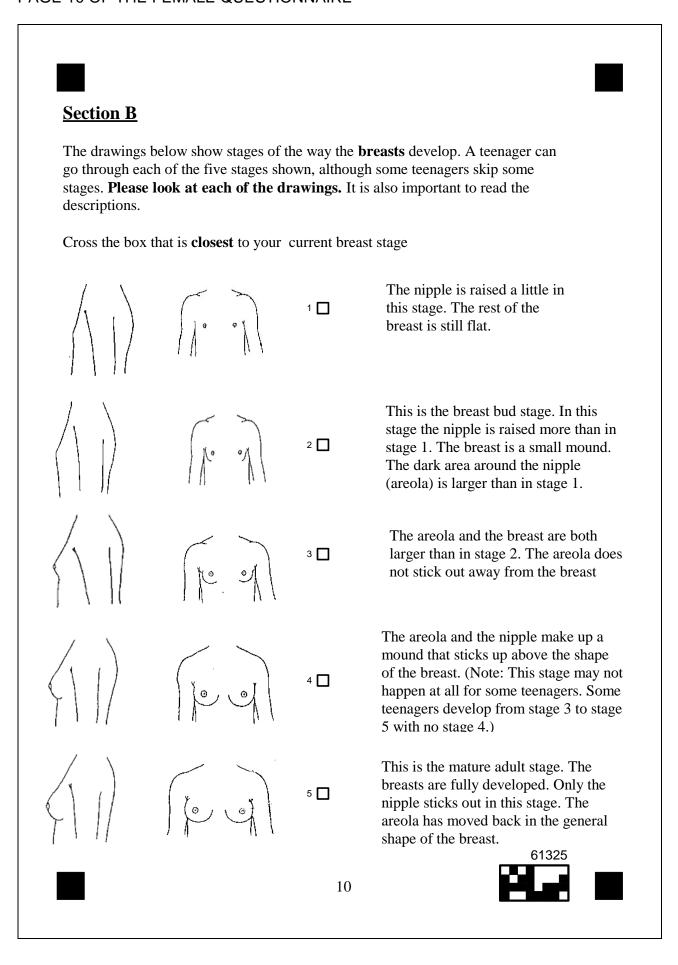
# pub 870 A11 (fem ale) / A5 (m ale): Hair has started to grow in responder armpits

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	1 Yes	4611	97.0	98.5	98.5
	2 No	72	1.5	1.5	100.0
	Total	4683	98.5	100.0	
Missing	-1 Not stated	72	1.5		
Total		4755	100.0		

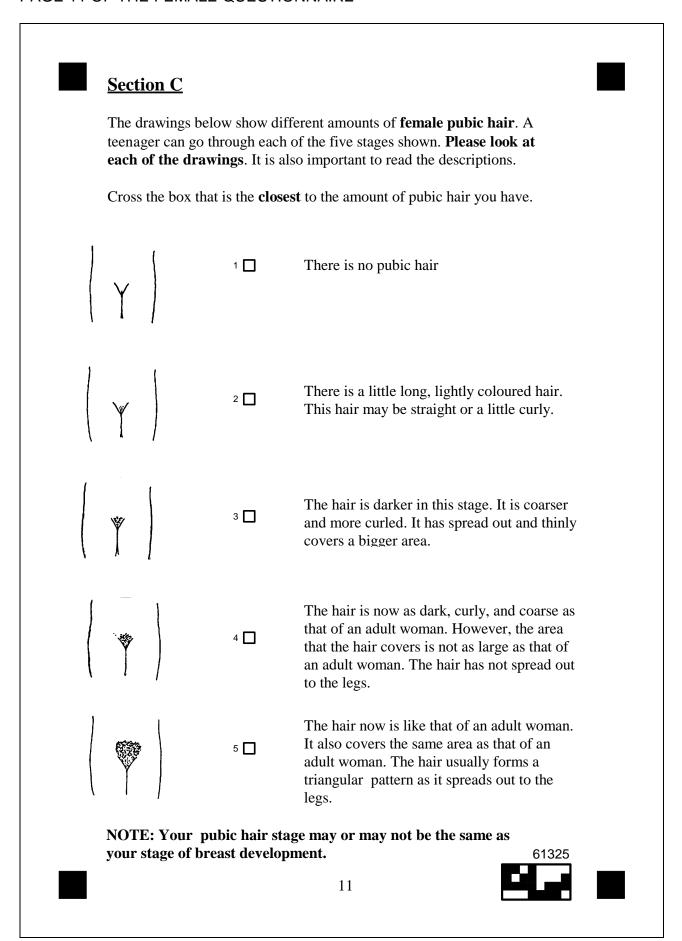
### PAGE 9 OF THE FEMALE QUESTIONNAIRE



### PAGE 10 OF THE FEMALE QUESTIONNAIRE



#### PAGE 11 OF THE FEMALE QUESTIONNAIRE



### PUB8 File – Sections B & C (female)

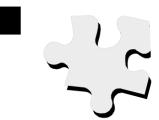
pub830 B (female): Development stage of breasts

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Stage 1	2	.0	.1	.1
	2 Stage 2	13	.3	.5	.5
	3 Stage 3	153	3.2	5.6	6.1
	4 Stage 4	1251	26.3	45.5	51.6
	5 Stage 5	1330	28.0	48.4	100.0
	Total	2749	57.8	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-1 Not stated	65	1.4		
	Total	2006	42.2		
Total		4755	100.0		

pub835  $\,$  C (fem ale): Development stage of pubic hair

	· , ,				
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Stage 1	16	.3	.6	.6
	2 Stage 2	5	.1	.2	.8
	3 Stage 3	69	1.5	2.5	3.3
	4 Stage 4	821	17.3	30.1	33.4
	5 Stage 5	1819	38.3	66.6	100.0
	Total	2730	57.4	100.0	
Missing	-7 Respondent is male	1941	40.8		
	-1 Not stated	84	1.8		
	Total	2025	42.6		
Total		4755	100.0		

### PAGE 5 OF THE MALE QUESTIONNAIRE



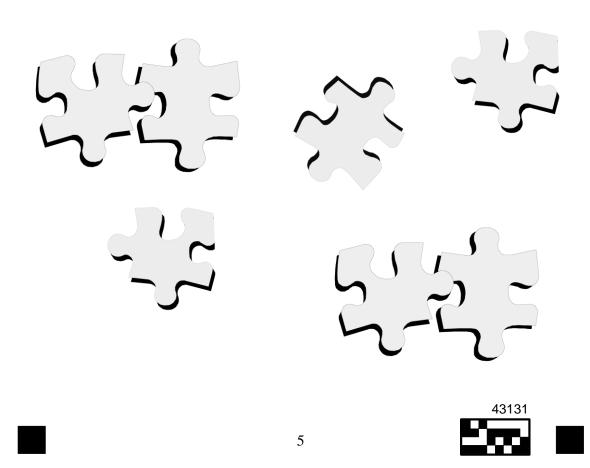
### PHYSICAL DEVELOPMENT

We would like to assess the stage of your physical development using the drawings on the next pages. These show various stages commonly used by doctors to assess the growth and development of boys.

We need to know which drawings most closely match your stage of development at the moment.

Not all teenagers follow the same pattern of development.

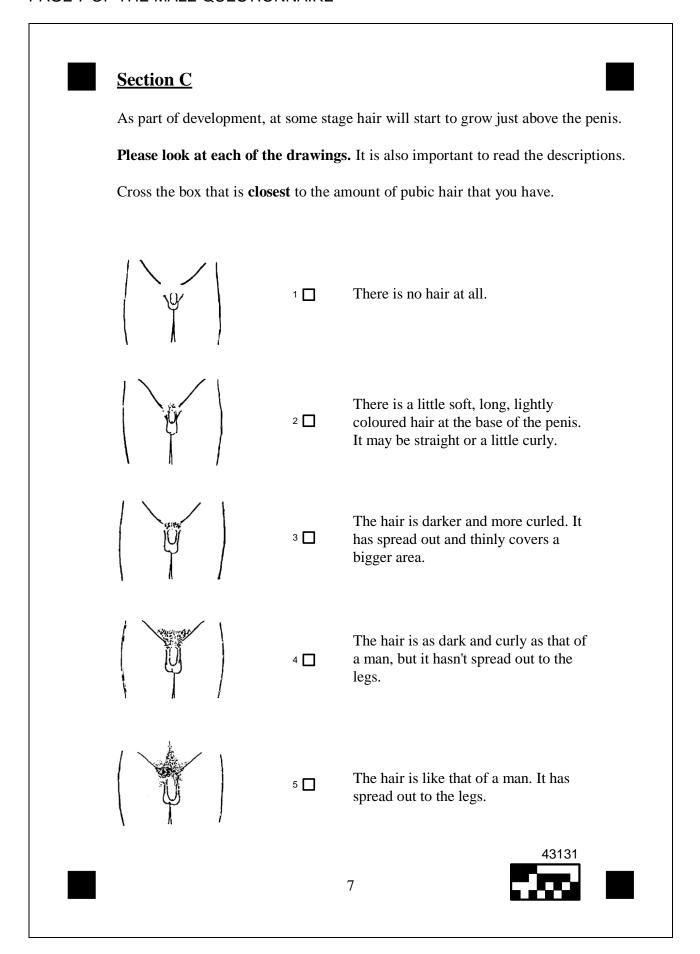
Just pick the stage that is closest, based on both the picture and the description.



### PAGE 6 OF THE MALE QUESTIONNAIRE

# **Section B** Teenagers go through the various stages of physical development at different ages. Some start as early as 6, others not until they are 20. We need your help in letting us know what stage you are at. **Please look at each of the drawings**. It is also important to read the descriptions. Cross the box that is **closest** to your current stage The size and shape of the testes, scrotum (the sac holding the testes) 1 🔲 and penis are about the same as when you were younger. The penis is a little bit bigger. The scrotum has dropped and the skin of 2 the scrotum has changed. The testes are bigger. The penis has grown longer, the 3 testes have grown and dropped lower. The penis is longer and wider. The 4 head of the penis is bigger, the scrotum is a darker colour and bigger. The testes are bigger. The penis, scrotum and testes are the 5 scrotum size and shape of a man's. 6

### PAGE 7 OF THE MALE QUESTIONNAIRE



## PUB8 File - Sections B & C (male)

pub850 B (male): Development stage of testes, scrotum and penis

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Stage 1	6	.1	.4	.4
	2 Stage 2	19	.4	1.1	1.5
	3 Stage 3	117	2.5	7.0	8.5
	4 Stage 4	612	12.9	36.6	45.1
	5 Stage 5	919	19.3	54.9	100.0
	Total	1673	35.2	100.0	
Missing	-7 Respondent is female	2814	59.2		
	-1 Not stated	268	5.6		
	Total	3082	64.8		
Total		4755	100.0		

pub855 C (male): Development stage of pubic hair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Stage 1	5	.1	.3	.3
	2 Stage 2	10	.2	.5	.8
	3 Stage 3	56	1.2	3.0	3.8
	4 Stage 4	614	12.9	32.5	36.3
	5 Stage 5	1202	25.3	63.7	100.0
	Total	1887	39.7	100.0	
Missing	-7 Respondent is female	2814	59.2		
	-1 Not stated	54	1.1		
	Total	2868	60.3		
Total		4755	100.0		

## PAGE 12 OF THE FEMALE QUESTIONNAIRE

Section D
D1. Did you have any help to fill this in?
No 1 🗆
Yes <sup>2</sup> □ ↓
If <u>yes</u> , please say who helped you:
a) A parent helped ¹□
b) Someone else helped ¹□
Day Month Year
D2. What is your date of birth? / 1 9 9
Day Month Year
D3. What is today's date?
Thank you VERY much for your help
When completed, please send this back to:
Children of the Nineties - ALSPAC 24 Tyndall Avenue Bristol BS8 1BR
coder ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ University of Bristol
http://www.alspac.bris.ac.uk/discovery
12

## PAGE 8 OF THE MALE QUESTIONNAIRE

Section D
D1. Did you have any help to fill this in?
No 1 🗆
Yes <sup>2</sup> □ ↓
If <u>yes</u> , please say who helped you:
a) A parent helped ¹□
b) Someone else helped ¹□
Day Month Year
D2. What is your date of birth? / 1 9 9
Day Month Year  Day / Day / Month / Year  Day / Day / Month / 2 0 0
Thank you VERY much for your help
When completed, please send this back to:
Children of the Nineties - ALSPAC 24 Tyndall Avenue Bristol BS8 1BR
coder Offfice use only   © University of Bristol
http://www.alspac.bris.ac.uk/discovery 43131
8

### **Question D1**

[Editing: If all three parts of question D1 were omitted then variables PUB890 to PUB892 were all set to -5. If either of the variables for parts a & b had value 1 then the initial indicator variable (PUB890) was set to 2 and values of -1 in PUB891 & PUB892 recoded to 2 'No'.]

pub890 D1: Respondent had help filling in the questionnaire

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No	4304	90.5	91.7	91.7
	2 Yes	391	8.2	8.3	100.0
	Total	4695	98.7	100.0	
Missing	-5 Question D1 omitted	60	1.3		
Total		4755	100.0		

pub891 D1a: A parent helped fill in the questionnaire

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	281	5.9	6.1	6.1
	2 No	4338	91.2	93.9	100.0
	Total	4619	97.1	100.0	
Missing	-5 Question D1 omitted	60	1.3		
	-1 Not stated	76	1.6		
	Total	136	2.9		
Total		4755	100.0		

pub892 D1b: Someone else helped fill in the questionnaire

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	36	.8	.8	.8
	2 No	4583	96.4	99.2	100.0
	Total	4619	97.1	100.0	
Missing	-5 Question D1 omitted	60	1.3		
	-1 Not stated	76	1.6		
	Total	136	2.9		
Total		4755	100.0		

### **Question D2**

[Editing: Note that this reported date of birth is not used.]

#### **Question D3**

[Editing: The date of completion was substituted with the date of receipt of the questionnaire from ALSPAC's administrative database if any of the following occurred: The date of completion was not fully completed, the date of completion was not a valid date (e.g. 31<sup>st</sup> November), the date of completion was later than the date of receipt or the date of completion was earlier than the date on which the questionnaire was sent out by ALSAPAC. This action was flagged in variable PUB896c. The month and year of completion were retained on the built file, but the day was dropped.]

pub896a D3: Date of completion of questionnaire - month

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 January	532	11.2	11.2	11.2
	2 February	230	4.8	4.8	16.0
	3 March	36	.8	.8	16.8
	4 April	18	.4	.4	17.2
	5 May	4	.1	.1	17.2
	6 June	19	.4	.4	17.6
	7 July	1401	29.5	29.5	47.1
	8 August	809	17.0	17.0	64.1
	9 September	582	12.2	12.2	76.4
	10 October	517	10.9	10.9	87.2
	11 November	437	9.2	9.2	96.4
	12 December	170	3.6	3.6	100.0
	Total	4755	100.0	100.0	

pub896b D3: Date of completion of questionnaire - year

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	2007	1775	37.3	37.3	37.3
	2008	2672	56.2	56.2	93.5
	2009	308	6.5	6.5	100.0
	Total	4755	100.0	100.0	

pub896c DV: Date of completion replaced with date of receipt

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes, incomplete d.o.c.	130	2.7	2.7	2.7
	2 Yes, illegal d.o.c.	2	.0	.0	2.8
	3 Yes, d.o.c. < d.o.s.	58	1.2	1.2	4.0
	4 Yes, d.o.c. > d.o.r.	45	.9	.9	4.9
	5 No	4520	95.1	95.1	100.0
	Total	4755	100.0	100.0	

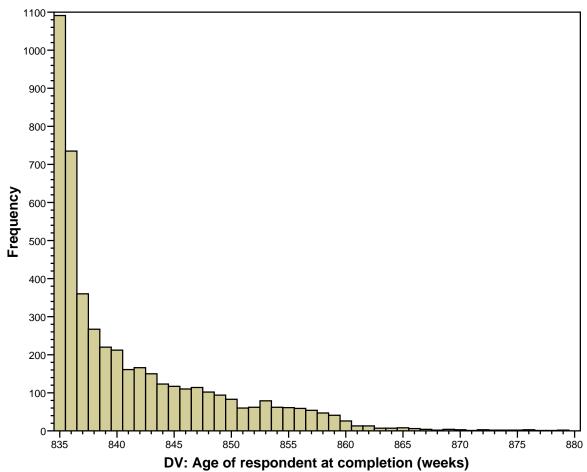
## **Derived Variables - Age at Completion**

The date of completion and the children's dates of birth from ALSPAC's central database, were used to calculate the child's age at completion in completed weeks and completed months.

pub897a DV: Age of respondent at completion (months)

		,			
		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	191	1	.0	.0	.0
	192	2510	52.8	52.8	52.8
	193	799	16.8	16.8	69.6
	194	512	10.8	10.8	80.4
	195	363	7.6	7.6	88.0
	196	284	6.0	6.0	94.0
	197	189	4.0	4.0	98.0
	198	44	.9	.9	98.9
	199	20	.4	.4	99.3
	200	10	.2	.2	99.5
	201	8	.2	.2	99.7
	202	4	.1	.1	99.8
	203	3	.1	.1	99.8
	204	1	.0	.0	99.9
	205	1	.0	.0	99.9
	207	1	.0	.0	99.9
	208	1	.0	.0	99.9
	209	1	.0	.0	99.9
	210	2	.0	.0	100.0
	211	1	.0	.0	100.0
	Total	4755	100.0	100.0	

pub897b DV: Age of respondent at completion (weeks)



plus the following missing values and outliers < 835 or > 880:

pub897b DV: Age of respondent at completion (weeks)

		Frequency	Percent	Valid Percent	Cumulativ e Percent
Valid	834	3	20.0	20.0	20.0
	881	1	6.7	6.7	26.7
	884	1	6.7	6.7	33.3
	885	1	6.7	6.7	40.0
	886	1	6.7	6.7	46.7
	887	1	6.7	6.7	53.3
	891	1	6.7	6.7	60.0
	902	1	6.7	6.7	66.7
	904	1	6.7	6.7	73.3
	910	1	6.7	6.7	80.0
	916	2	13.3	13.3	93.3
	919	1	6.7	6.7	100.0
	Total	15	100.0	100.0	

# **Appendix A**

This appendix contains the general coding instructions for scanned questionnaires that are referred to whenever any questionnaire is being coded. This document provides a general overview to the coding process as well as defining abbreviations for standard methods that are used in the specific coding instructions for the questionnaire (see Appendix B).

### PUB8 File - Appendix A

General coding instructions for *electronically scanned* ALSPAC questionnaires – first version, 5th December 2005.

These points are taken and adapted where necessary from the "General coding instructions for ALSPAC questionnaires – 3rd revision" document. They are to be used in conjunction with the specific instructions for each questionnaire.

#### **General guidelines:**

Check that the meaning of the question or the stated answer has not been adjusted by the respondent editing the question or qualifying the answer with words that make it not applicable.

Remember that a blank answer should always be left blank. It is not our task to infer what the respondent meant to say.

In cases where a box has been struck out, it should be considered as blank, meaning no answer given.

Logical editing will be taken care of by the data preparation team.

#### Dates

- a) If the <u>year is stated</u> but the day and/or month are left unfilled or are described as not known, then code the unknown element(s) as 99.
- b) In contrast, if the <u>year is blank or stated as not known</u>, but other elements of the date have been written, leave the unknown year *blank* so that there is no ambiguity with 1999.
- c) If the whole date is left totally blank, leave all blank.

#### **Coding instructions:**

Each type of questionnaire has its own unique coding sheet. To carry out the instructions by a standardised method, the abbreviations that have been used are:

- If more than one box is crossed, record the one with the lowest number in the specific range
- 8 If more than one box is crossed, record the one with the highest number in the specific range
- I If 'Yes' <u>and</u> 'no' are crossed in a <u>stand-alone</u> question, then leave the question blank. However, if the Yes/No have both been marked in a lead question with 'If Yes...' or 'If No...' after it, then the answer should be obvious from the information given after the 'If Yes/No'.

#### II Code as a number.

Make sure that the answer is in the required units, e.g. weeks rather than months, or pints rather than glasses, or whatever. There will be rules in the coding instructions as to whether to round fractions up or down, for each question where it is likely to occur.

Other possible 'non-standard' indications by the respondent:

If answer stated as not known, code as 9, 99, 999 or 9999 - depending on field length. For this occurrence with <u>dates</u> see above.

If "occasional" is stated where a number is required, this is usually coded as 97. There are occasionally other codes in the 90 range to be used as indicated in the coding instructions.

If "none" or "nil" is stated for a numeric response, code as zero.

If there is no response, or they reply "Not applicable", "N/A" or something equivalent, leave the boxes blank.

IV Code as ddmmyy for days, months, years. See instructions above about dates.

## PUB8 File - Appendix A

### "Other, please describe" questions

Where information is written in an "other, please describe" box, check various points:

- 1. The information given there should not fit into one of the <u>other categories</u> in the question. If it is the same, recode it as that category, but if there is doubt, or it is clearly different, leave it as "other".
- 2. If there is information written in the box, it should have the corresponding box marked. If it is not possible to assign an answer to a corresponding box, (which happens often in qualitative answers) then assign the code "zero" indicating a relevant comment has been given, but we don't know which "yes" category to put it into.

# **Appendix B**

This appendix contains the coding instructions specific to this questionnaire. Note that there are separate instructions for the male and female versions. The instructions for the female version are presented first.

### PUB8 File - Appendix B

<u>Coding instructions for: sixth version (03/03/06) "Growing and Changing (6) – Girl Teenager Questionnaire"</u> (fomerly known in Versions 1-5 as "Mother/Daughter Growing and Changing")

These questionnaires have all been electronically scanned.

### Coding the "Growing and Changing (6) - Girl Teenager Questionnaire " replies

For data preparation follow the general rules outlined in the "General Coding Instructions For Electronically Scanned ALSPAC Questionnaires – First Version, 5th December 2005".

# <u>In addition, run through the whole questionnaire to make sure that any names or other identifiers of the</u> respondent are deleted. These are not always in the text answers.

#### **Front Cover**

Check that the questionnaire number is present, and that the date of form is 03/03/06

Page 4, Section A		
A1	II, II	Round fractions down. If the height is given in more than one form, leave them both in. Convert 12+ inches to feet and inches. Convert 100+ cms to metres and centimetres.  If they have written NK or equivalent, code with 9s throughout.  Note that quite unrealistic and impossible answers are sometimes written in this question when respondents cannot cope with imperial and metric equivalents.
A2	II, II	As A1, but converting 'pounds' to 'stones & pounds' if necessary ( <b>14lbs</b> = 1 stone!). Again, note that quite unrealistic and impossible answers are sometimes written in this question when respondents cannot cope with imperial and metric equivalents.
A3	7	
Page 5		
A4		If Yes and No are both crossed, code as blank unless part a) is answered, in which case code as Yes.  If No is answered, ensure the skip to A10.
A4a	II	Round age down to completed years.
A5	II	Code as month, year. Code 'Not Known' or equivalent as 99 9999. If not yet started, should be blank.
A6a	II	If not yet started, should be blank. If 'It varies' is crossed as well as 'number of days', delete the number.
A6b	8	
A7	II	If not yet started, should be blank. If a range is given, code to the midpoint. If 'It varied' is crossed as well as 'number of days', delete the number.
Page 6		
A8a		If Yes and No are both crossed, code as blank unless there is a meaningful answer in part i), in which case code as Yes.  If No is answered, ensure the skip to A8b

## PUB8 File – Appendix B

A8ai		If Yes and No are both crossed, code as blank.
A8b		If Yes and No are both crossed, code as blank unless there is a meaningful answer in part i), in which case code as Yes.
A8bi		If Yes and No are both crossed, code as blank.
A8c		If Yes and No are both crossed, code as blank unless there is a meaningful answer in part i), in which case code as Yes.
A8ci		If Yes and No are both crossed, code as blank.
Page 7		
A9	8	
A10a	8	
A10b		If a relevant description is given in A10b, please ensure that the "yes" box is crossed in A10a. Please note that the text response is not being coded here.
A11		If Yes and No are both crossed, code as blank.
Page 10, Section B		
Breast development	8	
Page 11, Section C		
Hair growth	8	
Page 12, Section D		
D1		If both "Yes" and "No" are answered, code as blank <u>unless</u> an answer has been given in a) or b) in which case code as 2.
D1a,b		There should be either a cross or a blank here.
D2	IV	The year should be either 91, 92 or Jan/Feb (only) 93, as there were no female cases born in 1990.
D3	IV	If this question is blank, copy the date of receipt from the front cover. Watch out for dates at the turn of the year as the stated year is sometimes incorrect!

## END OF QUESTIONNAIRE

Growing & Changing (6) – Girl Teenager Questionnaire (03/03/06)

Sheila Preece 19/04/06

### PUB8 File - Appendix B

# <u>Coding instructions for: sixth version (02/03/06) "Growing and Changing (6) – Male Teenager Questionnaire"</u> (fomerly known in Versions 1-5 as "Parent/Son Growing and Changing")

These questionnaires have all been electronically scanned.

### Coding the "Growing and Changing (6) - Male Teenager Questionnaire " replies

For data preparation follow the general rules outlined in the "General Coding Instructions For Electronically Scanned ALSPAC Questionnaires – First Version, 5th December 2005".

# <u>In addition, run through the whole questionnaire to make sure that any names or other identifiers of the respondent are deleted. These are not always in the text answers.</u>

#### **Front Cover**

Check that the questionnaire number is present, and that the date of form is 02/03/06

Page 4, Section A		
A1	II, II	Round fractions down. If the height is given in more than one form, leave them both in. Convert 12+ inches to feet and inches. Convert 100+ cms to metres and centimetres.  If they have written NK or equivalent, code with 9s throughout.  Note that quite unrealistic and impossible answers are sometimes written in this question when respondents cannot cope with imperial and metric equivalents.
A2	II, II	As A1, but converting 'pounds' to 'stones & pounds' if necessary ( <b>14lbs</b> = 1 stone!). Again, note that quite unrealistic and impossible answers are sometimes written in this question when respondents cannot cope with imperial and metric equivalents.
A3	7	
A4	8	
A5		If Yes and No are both crossed, code as blank.
Page 6, Section B		
Penis development	8	
Page 7, Section C		
Hair growth	8	
Page 8, Section D		
D1		If both "Yes" and "No" are answered, code as blank <u>unless</u> an answer has been given in a) or b) in which case code as 2.
D1a,b		There should be either a cross or a blank here.

## PUB8 File – Appendix B

D2	IV	There is only one male case in the cohort born in 1990. The year should otherwise be either 91, 92 or Jan/Feb (only) 93.
D3	IV	If this question is blank, copy the date of receipt from the front cover. Watch out for dates at the turn of the year as the stated year is sometimes incorrect!

## END OF QUESTIONNAIRE

Growing & Changing (6) – Male Teenager Questionnaire (02/03/06)

Sheila Preece 19/04/06

coding specifications\nc06\chG&C Boy

## **Appendix C: Questionnaire Methodology**

### Administration

Questionnaires were identified by a 9 digit check-summed identifier (QUESTIONNAIRE ID) that is unique to both the recipient and the particular questionnaire. These numbers, together with bar codes, were printed on sticky labels that were then applied to the front pages of the questionnaires prior to being mailed out. If a response was not received after 3 weeks a reminder letter was sent.

The receipt of questionnaires was logged by the Questionnaire Administration Department using a bar code scanning system, which recorded the date of receipt. If returned blank by the respondent (but not, for example, if returned as "not known at this address") this was flagged. The questionnaires were examined and if badly completed (e.g. marks not in boxes) or if text was written outside of supplied boxes then the office use box on the back page was marked in order to flag that detailed checking was required at the verification stage (see below). Any signed comments on the back page were dealt with as appropriate and the completed questionnaires batched together in batches of 20, separately for the male and female versions.

### **Data handling**

The questionnaires were designed using the Teleform data capture software. They were the first questionnaires completed by the study children to be designed in this way, although the Teleform system had previously been used for the carer and partner adult learning questionnaires and for staff completed forms from the Focus clinics. The questionnaires were booklets with A5 pages, stapled in the folded spine. The expected responses to the questions fall into four categories: self-coding tick boxes, numeric, free text and occasionally constrained print (text entered one letter to a box).

Completed questionnaires had the spine guillotined off and were then fed into a document scanner. The data were captured from the resulting images by the Teleform software. Any anomalies, such as multiple ticks or numeric / constrained print responses that the system could not interpret were flagged by the software and corrected by an operator at the verification stage. During verification the captured free text was also checked for accuracy. The rules for correcting multiple ticks and other anomalies were set out in the ALSPAC general coding instructions (see Appendix A) and the specific coding instructions for this questionnaire (see Appendix B). The numeric, tick box and any constrained print data were then exported to an SPSS data file and the free text exported to a database. The images of the forms were exported to the ADOS image storage system for archiving.

The SPSS data files then had appropriate variable and value labels applied. The administrative identifier (QUESTIONNAIRE ID) was converted to the research identifiers ALN & QLET. At the end of this process the data were classified as clean; they are deemed to accurately represent what the respondent actually entered on the questionnaire, except that the text responses have not been processed. The free text is available on request under special confidentiality rules for projects that require it.

In order to prepare the data for general release the clean data were matched to the information about dispatch and return held in the ALSPAC administrative database. This enabled creation of variables for all members of the cohort indicating whether or not a

### PUB8 File - Appendix C

questionnaire was sent out, whether it was returned etc. The variables were renamed according to a unique system to ensure there are no conflicts with variable on other data files. All variable names start with the letters "pub" which are followed by three digits (the first of which is always an 8) and occasionally a further letter. The data were then edited, adjusting for illogical responses, dealing with skip statements and adding useful derived variables, as described in the relevant sections of this documentation.