

Appendix B. Weighting methods in DNSIYC

Sarah Tipping

Table of contents

B.1	Background.....	1
B.2.	Interview weighting factors	1
B.3.	Clinic weighting factors.....	3
B.4.	Blood sample weighting factors.....	4
B.5.	Urine sample weighting factors.....	4

List of tables

Table B.1	6
Table B.2	7
Table B.3	8
Table B.4	10
Table B.5	11
Table B.6	12
Table B.7	15

List of figures

Figure B.1. Overview of methods used for the interview weighting factor . 3

B.1 Background

The Diet and Nutrition Survey of Infants and Young Children (DNSIYC) required a set of weighting factors to adjust the sample for differences in sample selection and response. The weighting factors adjust for differential selection probabilities of boost sample members, non-response to the individual questionnaire, non-response to the clinic visit and non-response to providing a blood sample.

B.2. Interview weighting factors

An interview weighting factor was required for the 2,683 parents who responded to the individual interview and completed three or four food diary days for their child. This weighting factor was generated using a combination of logistic regression modelling and calibration. The aim was to reduce bias resulting from sampling error and differential non-response.

The first step was to model response behaviour to the interview using logistic regression. The DNSIYC sample contained two boost samples; a boost of Healthy Start recipients and a boost sample of individuals living in Scotland. The modelling was therefore carried out separately for the core sample from England, Wales and Northern Ireland and for the core and boost sample from Scotland. The Healthy Start recipients were excluded at this step. The weighting factors from the Scottish non-response model fed into both the overall weighting factors and into a separate weighting factor for analysis of the Scottish sample only. Having a separate model for Scotland ensures any estimates for Scotland in the overall sample match those produced for the Scottish-only sample.

A logistic regression models the relationship between an outcome variable (response to the interview) and a set of predictor variables. The predictor variables were a set of address, household and individual characteristics taken from the Child Benefit register. The model was used to generate predicted probabilities of response. For each participant, this is the predicted probability of them taking part in the interview, given their individual characteristics and the characteristics of their household. Participants with characteristics associated with non-response receive a low predicted probability. The predicted probabilities were then used to generate a set of non-response weighting factors; a larger weighting factor was applied to participants with a low predicted probability, increasing their representation in the sample. The full non-response models for core England, Wales and Northern Ireland and for Scotland (core plus boost) are given in Tables B.1 and B.2.

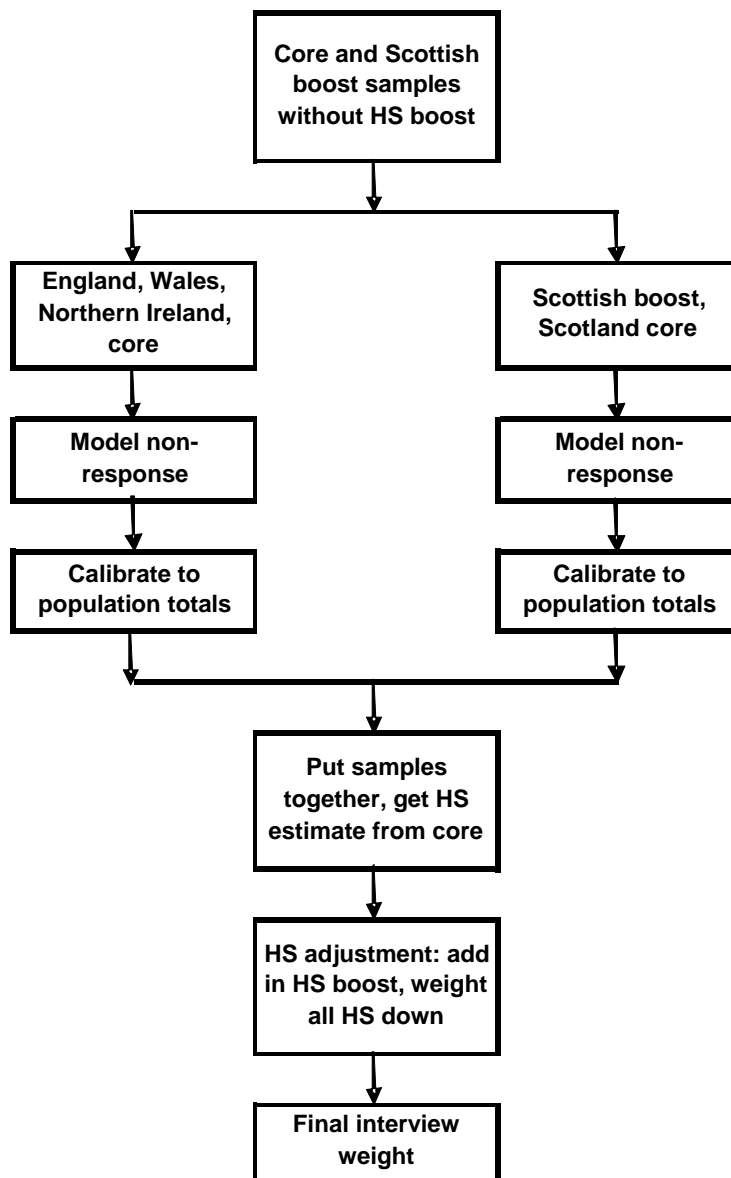
Tables B.1 and B.2

The next step was to calibrate the non-response weighting factors. Again, this was carried out separately for core England, Wales and Northern Ireland samples and for the combined Scottish core and boost samples. An iterative procedure was used to adjust the non-response weighting factor until the distribution of the (weighted) sample matched that of the population for a set of key variables. The adjustment keeps the values of the final weighting factors as close as possible to those of the initial weighting factors, which ensures the properties of the initial non-response weighting factors are retained in the final calibrated weights. This step made the profile of the sample match the population for child's age and gender, age of mother at the time of the child's birth and region. The population figures for calibration were taken from birth counts¹. The profiles of the population and the final weighted sample are shown in Table B3.

Table B.3

The calibrated weighting factors were combined into a single weighting factor. All core UK and Scottish boost cases had a calibrated non-response weighting factor. The final stage of the interview weighting factors was to incorporate the Healthy Start boost into the sample. The core UK and Scottish boost data (excluding the Healthy Start boost) was used to get a weighted estimate of the proportion of Healthy Start recipients. It was found that 20.6% of the weighted core and Scottish boost sample were in receipt of Healthy Start vouchers. The Healthy Start boost recipients were then added into the sample, increasing the proportion of Healthy Start recipients to 23.3%. All Healthy Start recipients were therefore weighted down by a factor of 20.6/23.3, so that the combined proportion of Healthy Start recipients matched the weighted estimate from the core data. This weighting factor was scaled, meaning the weighted sample size matched the unweighted sample size and the weights had a mean of one. This scaled weighting factor is the final interview weight. Figure B.1 gives an overview of this process.

Figure B.1. Overview of methods used for the interview weighting factor



B.3. Clinic weighting factors

DNSIYC also required a weighting factor for non-response to the clinic visit and a weighting factor for non-response to the blood sample.

All core and Healthy Start parents were invited to attend a clinic where anthropometric measurements could be taken. Scottish boost sample members were not eligible. The first step of the clinic weighting factors was to generate a clinic-specific interview weighting factor, i.e. an interview weighting factor for all cases that were eligible for the clinic visit. This second interview weighting factor was generated using the same methods as before but only the core sample was included (the Scottish boost was excluded). The same variables were included in

the response model and the non-response weighting factors were adjusted in a single stage to UK totals. The sole purpose of this weighting factor was to feed into the clinic weighting factors.

The second step of the clinic weighting factors was to model non-response to the clinic visit. 2,212 participants were eligible for the clinic visit, 973 attended. Response behaviour to the clinic visit was modelled using a logistic regression model. The outcome variable was whether or not an eligible individual attended a clinic and the predictor variables were taken from the DNSIYC interview. The full model is shown in Table B.4. The predicted probabilities of response produced by the model were used to generate clinic non-response weighting factors. These non-response weighting factors were combined with the clinic-specific interview weighting factors to give the final clinic weighting factors. The weighting factors were scaled, so the mean weighting factor equalled one and the weighted sample size matched the unweighted sample size.

Table B.4

The clinic weighting factors should be used to analyse any data collected during the clinic visit, with the exception of blood sample data. Blood sample data has a separate weighting factor due to the higher levels of non-response to blood sample collection.

B.4. Blood sample weighting factors

Participants who attended a clinic were asked if their child would give a blood sample. Five hundred and thirteen usable samples were collected from participating children. As before, non-response to the blood sample was modelled using a logistic regression model. Information collected at the interview and clinic visit was used to model response to the blood sample. The full non-response model is given in Table B.5. The non-response weighting factors from the model were combined with the final clinic weighting factors to give the final blood weighting factors. These weighting factors correct for non-response to the individual questionnaire, the clinic visit and the blood sample and should be used for any analysis of blood sample data. The weighting factors were scaled, so the mean weighting factor equalled one and the weighted sample size matched the unweighted sample size. The profiles of the weighted clinic and blood samples are shown in Table B6.

Tables B.5 and B.6

B.5. Urine sample weighting factors

As well as being asked for a blood sample, the participants who attended a clinic were also asked if a urine sample could be taken from their child. A total of 622

usable samples were collected from participating children. The methods used to create weighting factors to compensate for non-response to the urine sample were the same as those used to create weighting factors for non-response to the blood sample; logistic regression was used to model non-response behaviour using information collected at both the interview and clinic visit. The full non-response model is given in Table B.7. The non-response weighting factors from the model were combined with the final clinic weighting factors to give the final urine weighting factors. These weighting factors correct for non-response to the individual questionnaire, the clinic visit and the urine sample and should be used for any analysis of urine sample data. The weighting factors were scaled, so the mean weighting factor equalled one and the weighted sample size matched the unweighted sample size. The profiles of the weighted clinic and urine samples are shown in Table B.6.

Table B.7

References and endnotes

¹ This means the weights also account for any non-take up of Child Benefit, although it can be seen in Table 3 that the distributions were very similar for each.

Table B.1

Non-response model for DNSIYC interview weighting factors – core England, Wales and NI

	B	S.E.	Wald	df	Sig.	Exp(B)
Wave of sampling	0.16	0.075	4.7	1	0.030	1.18
Recipient's title			14.9	3	0.002	
Miss					(baseline)	
Mr	-0.35	0.150	5.3	1	0.021	0.71
Mrs	0.17	0.091	3.5	1	0.062	1.19
Ms	0.12	0.199	0.3	1	0.561	1.12
Recipient's age in years at sampling			24.8	4	0.000	
<25					(baseline)	
25-29	0.40	0.113	12.6	1	0.000	1.49
30-34	0.50	0.119	17.4	1	0.000	1.64
35-39	0.62	0.135	21.0	1	0.000	1.86
40+	0.54	0.184	8.6	1	0.003	1.71
Government Office Region			26.0	10	0.004	
North East					(baseline)	
North West	-0.14	0.214	0.4	1	0.508	0.87
Yorkshire and The Humber	-0.02	0.226	0.0	1	0.937	0.98
East Midlands	-0.02	0.234	0.0	1	0.950	0.99
West Midlands	0.02	0.222	0.0	1	0.922	1.02
East of England	-0.16	0.222	0.5	1	0.462	0.85
London	-0.57	0.215	7.0	1	0.008	0.57
South East	-0.32	0.210	2.3	1	0.132	0.73
South West	-0.04	0.229	0.0	1	0.867	0.96
Northern Ireland	0.26	0.285	0.9	1	0.354	1.30
Wales	-0.07	0.257	0.1	1	0.794	0.94
Number of children aged 0-16 yrs in household			14.7	4	0.005	
1					(baseline)	
2	-0.04	0.087	0.2	1	0.682	0.97
3	-0.25	0.120	4.4	1	0.036	0.78
4	-0.40	0.186	4.6	1	0.031	0.67
5+	-0.79	0.279	8.0	1	0.005	0.45
Population density - quintiles			18.2	4	0.001	
1 - least dense					(baseline)	
2	0.08	0.131	0.3	1	0.556	1.08
3	-0.04	0.122	0.1	1	0.748	0.96
4	-0.28	0.117	5.9	1	0.016	0.75
5 - most dense	-0.42	0.139	8.9	1	0.003	0.66
Age in months at sampling			18.7	6	0.005	
4-5 mth					(baseline)	
6-7 mth	-0.35	0.175	4.1	1	0.043	0.70
8-9 mth	-0.36	0.173	4.4	1	0.036	0.70
10-11 mth	-0.23	0.173	1.7	1	0.194	0.80
12-13 mth	-0.53	0.174	9.4	1	0.002	0.59
14-15 mth	-0.27	0.175	2.5	1	0.116	0.76
16-17 mth	-0.66	0.190	11.9	1	0.001	0.52
Constant	0.71	0.257	7.6	1	0.006	2.03

The response is 1 = individual responded to the interview, 0 = non response

Only variables that are significant at the 0.05 level are included in the model

The model R2 is 0.048 (Cox and Snell)

B is the estimate coefficient with standard error **S.E.**

The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom *df*.

If the test is significant (sig. < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model

Table B.2

Non-response model for DNSIYC interview weighting factors – Scotland core and boost

	B	S.E.	Wald	df	Sig.	Exp(B)
Wave of sampling	-0.13	0.146	0.8	1	0.379	0.88
Recipient's title			4.2	3	0.241	
Miss					(baseline)	
Mr	-0.25	0.302	0.7	1	0.401	0.78
Mrs	0.24	0.172	1.9	1	0.166	1.27
Ms	-0.11	0.357	0.1	1	0.749	0.89
Recipient's age in years at sampling			20.3	4	0.000	
<25					(baseline)	
25-29	0.39	0.206	3.5	1	0.062	1.47
30-34	0.79	0.214	13.7	1	0.000	2.21
35-39	0.91	0.263	11.9	1	0.001	2.48
40+	1.27	0.382	11.0	1	0.001	3.56
Number of children aged 0-16 yrs in household			0.2	2	0.913	
1					(baseline)	
2	0.07	0.160	0.2	1	0.670	1.07
3+	0.04	0.219	0.0	1	0.874	1.04
Population density - quintiles			5.0	4	0.285	
1 - least dense					(baseline)	
2	0.46	0.230	4.0	1	0.045	1.58
3	0.09	0.187	0.2	1	0.635	1.09
4	-0.07	0.228	0.1	1	0.761	0.93
5 - most dense	0.00	0.231	0.0	1	0.987	1.00
Age in months at sampling			2.2	6	0.897	
4-5 mth					(baseline)	
6-7 mth	-0.28	0.325	0.8	1	0.384	0.75
8-9 mth	-0.37	0.322	1.3	1	0.253	0.69
10-11 mth	-0.41	0.325	1.6	1	0.205	0.66
12-13 mth	-0.23	0.322	0.5	1	0.485	0.80
14-15 mth	-0.26	0.321	0.6	1	0.424	0.77
16-17 mth	-0.41	0.364	1.3	1	0.261	0.67
Constant	0.25	0.323	0.6	1	0.447	1.28

The response is 1 = individual responded to the interview, 0 = non response

Only variables that are significant at the 0.05 level are included in the model

The model R2 is 0.049 (Cox and Snell)

B is the estimate coefficient with standard error **S.E.**

The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom *df*.

If the test is significant (sig. < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model

Table B.3

Distribution of the weighted interview sample, issued sample and population figures

	Population	Achieved sample1	Selected sample2
		Weighted by interview weighting factor	Scottish boost selection weighting factors only
	%	%	%
GOR			
North East	4	4	4
North West	11	11	11
Yorks and Humber	8	8	8
East Midlands	7	7	7
West Midlands	9	9	9
East of England	9	9	9
London	17	17	15
South East	13	13	14
South West	8	8	8
Northern Ireland	3	3	4
Scotland	7	7	8
Wales	4	4	5
IMD quintiles (from sampling frame)			
1 - least deprived	20	21	22
2	20	18	19
3	20	18	18
4	20	25	23
5 - most deprived	20	18	18
Population density quintiles (from sampling frame)			
1 - least dense	20	23	23
2	20	16	16
3	20	20	20
4	20	21	21
5 - most dense	20	21	20
Recipient's age - grouped (HMRC population counts)			
<25	22	22	22
25-29	25	25	25
30-34	28	28	28
35-39	18	18	19
40+	6	6	7
Recipient's gender (HMRC population counts)			
Male	10	11	10
Female	91	89	90
Mother's age at child's birth (birth records)			
<25	25	25	Not known for selected sample
25-29	28	28	
30-34	28	28	
35-39	16	16	

40+	4	4	
Child's gender (HMRC population counts)			
Male	51	51	50
Female	49	49	50
Child's age at sampling (HMRC population counts)			
0-2 months	16	16	16
3-4 months	16	17	17
5-6 months	16	16	17
7-8 months	17	17	18
9-10 months	18	17	16
11-12 months	17	17	17
Child's gender (birth records)			
Male	51	51	50
Female	49	49	50
Total number of children in household (HMRC population counts)			
1	49	51	50
2	33	32	32
3	13	12	13
4+	6	6	6
<i>Unweighted bases</i>	<i>748,480</i>	<i>2,586</i>	<i>4,276</i>

¹Healthy Start boost not included, Scottish boost included but weighted down

²Healthy Start and Scottish boosts included

Table B.4

Non-response model for DNSIYC clinic weighting factors

	B	S.E.	Wald	df	Sig.	Exp(B)
Wave	0.10	0.088	1.2	1	0.273	1.10
HSVou	-0.14	0.126	1.2	1	0.284	0.87
Mothers age at birth	0.01	0.008	3.2	1	0.074	1.01
Child's age at interview			7.6	3	0.054	
4-6 months					(baseline)	
7-9 months	-0.25	0.155	2.7	1	0.103	0.78
10-12 months	-0.36	0.156	5.2	1	0.023	0.70
13-18 months	-0.39	0.148	6.9	1	0.008	0.68
Child's gender			0.5	1	0.486	
Male					(baseline)	
Female	-0.06	0.088	0.5	1	0.486	0.94
Government Office Region			20.8	11	0.035	
East Midlands					(baseline)	
East of England	-0.01	0.221	0.0	1	0.973	0.99
London	-0.20	0.219	0.8	1	0.361	0.82
North East	0.13	0.275	0.2	1	0.629	1.14
North West	-0.23	0.214	1.2	1	0.274	0.79
Northern Ireland	-0.10	0.307	0.1	1	0.741	0.90
Scotland	0.25	0.232	1.1	1	0.291	1.28
South East	0.13	0.204	0.4	1	0.527	1.14
South West	0.36	0.230	2.5	1	0.115	1.44
Wales	-0.25	0.273	0.9	1	0.352	0.78
West Midlands	0.32	0.220	2.1	1	0.149	1.37
Yorkshire and The Humber	-0.16	0.228	0.5	1	0.482	0.85
Population density - quintiles			7.9	4	0.096	
1 - least dense					(baseline)	
2	-0.18	0.149	1.4	1	0.239	0.84
3	-0.35	0.137	6.5	1	0.011	0.70
4	-0.25	0.137	3.3	1	0.069	0.78
5 - most dense	-0.08	0.161	0.3	1	0.606	0.92
Respondent's highest qualification			13.6	2	0.001	
None					(baseline)	
Degree level or above,	0.62	0.169	13.6	1	0.000	1.86
Below degree level	0.46	0.152	9.1	1	0.003	1.58
Number of adults aged 16 and over in household			1.2	2	0.553	
1					(baseline)	
2	-0.04	0.139	0.1	1	0.762	0.96
3+	-0.20	0.192	1.1	1	0.305	0.82
Constant	-0.57	0.348	2.7	1	0.100	0.56

The response is 1 = individual responded to the clinic, 0 = non response to the clinic visit

Only variables that are significant at the 0.05 level are included in the model

The model R2 is 0.027 (Cox and Snell)

B is the estimate coefficient with standard error **S.E.**

The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom *df*.

If the test is significant (*sig.* < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model

Table B.5

**Non-response model for DNSIYC
blood sample weighting factors**

	B	S.E.	Wald	df	Sig.	Exp(B)
Wave of sampling	0.08	0.138	0.3	1	0.568	1.08
In receipt of HS vouchers	-0.09	0.195	0.2	1	0.633	0.91
Mothers age in years	0.01	0.013	0.2	1	0.648	1.01
Child's age at interview			2.9	3	0.404	
4-6 months					(baseline)	
7-9 months	0.34	0.245	1.9	1	0.171	1.40
10-12 months	0.39	0.245	2.5	1	0.113	1.47
13-18 months	0.37	0.231	2.5	1	0.113	1.44
Child's gender			0.0	1	0.935	
Male					(baseline)	
Female	-0.01	0.138	0.0	1	0.935	0.99
Government Office Region			52.6	11	0.000	
East Midlands					(baseline)	
East of England	-0.32	0.331	1.0	1	0.327	0.72
London	1.22	0.320	14.5	1	0.000	3.38
North East	0.20	0.418	0.2	1	0.632	1.22
North West	-0.31	0.317	0.9	1	0.331	0.74
Northern Ireland	0.11	0.457	0.1	1	0.805	1.12
Scotland	-0.77	0.354	4.7	1	0.031	0.47
South East	0.13	0.309	0.2	1	0.672	1.14
South West	-0.25	0.344	0.5	1	0.474	0.78
Wales	0.00	0.403	0.0	1	0.996	1.00
West Midlands	0.15	0.336	0.2	1	0.651	1.16
Yorkshire and The Humber	0.19	0.339	0.3	1	0.583	1.21
Household size			12.5	4	0.014	
2					(baseline)	
3	0.11	0.306	0.1	1	0.713	1.12
4	-0.23	0.313	0.5	1	0.465	0.80
5	0.52	0.341	2.3	1	0.129	1.68
6+	-0.08	0.373	0.0	1	0.830	0.92
Tenure			4.9	2	0.087	
Owner occupier (inc shared ownership)					(baseline)	
Rent - Housing Association	0.09	0.209	0.2	1	0.680	1.09
Rent - Private landlord	0.39	0.183	4.6	1	0.033	1.48
Mother's height measurement taken			6.9	1	0.009	
No					(baseline)	
Yes	0.79	0.300	6.9	1	0.009	2.20
Constant	-1.28	0.652	3.9	1	0.049	0.28

The response is 1 = child gave a usable blood sample, 0 = child did not give a usable blood sample

Only variables that are significant at the 0.05 level are included in the model

The model R2 is 0.095 (Cox and Snell)

B is the estimate coefficient with standard error **S.E.**

The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df.

If the test is significant (sig. < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model

Table B.6

Distribution of the weighted clinic and blood samples

	Eligible for clinic Weighted by clinic-specific ¹ interview weighting factor %	Attended clinic (eligible for blood and urine samples) Weighted by clinic weighting factor %	Achieved blood sample Weighted by blood weighting factor %	Achieved urine sample Weighted by urine weighting factor %
Government Office Region				
East Midlands	7	7	7	7
East of England	9	9	9	9
London	17	16	16	17
North East	4	4	4	4
North West	11	11	11	11
Northern Ireland	3	3	3	3
Scotland	7	7	8	7
South East	13	13	13	13
South West	8	8	8	8
Wales	4	4	5	4
West Midlands	9	9	9	9
Yorkshire and The Humber	8	8	8	8
Mothers age at interview				
<25 years	22	21	20	20
25-29 years	25	27	28	27
30-34 years	29	28	29	29
35-39 years	18	17	18	18
40+ years	6	7	7	7
Household size				
2	6	6	7	6
3	40	40	39	40
4	30	31	31	29
5	15	15	15	15
6+	9	8	8	8
Number of adults aged 16 and over in household				
1	14	15	15	14
2	76	77	76	77
3+	9	9	8	9
Child's age at interview				
4-6 months	11	11	11	11
7-9 months	26	25	25	26
10-12 months	25	25	24	25
13-18 months	38	39	40	38
Child's gender				
Male	51	52	51	52

Female	49	48	49	48
Child's ethnic group				
Not white	81	82	83	82
White	19	18	17	18
UK Child's National Identity				
English	45	44	43	43
Scottish	5	5	4	5
Welsh	3	3	3	2
Irish	1	1	2	1
British	38	38	37	39
Other	9	10	11	10
HRP work status				
Never	5	5	5	5
Yes, now	77	77	76	77
Yes, have worked in the past	19	19	19	18
Mother's work status				
Studying	2	2	1	1
Work FT/PT	45	45	46	45
Not in employment	53	54	53	54
HRP work FT or PT				
Not working	5	5	5	5
Full time	74	76	78	76
part time	22	19	18	19
NS-SEC				
1 Managerial and professional occupations	40	41	41	43
2 Intermediate occupations	10	10	11	9
3 Small employers and own account workers	9	9	9	8
4 Lower supervisory and technical occupations	7	7	7	8
5 Semi-routine and routine occupations	29	28	26	27
6 Not classifiable	6	5	6	6
Tenure				
Owner occupier (inc shared ownership)	53	53	54	55
Rent - Housing Association	20	21	21	19
Rent - Private landlord	27	26	26	27
In receipt of HS vouchers				
Yes	20	20	20	20
No	80	80	80	80
Mother's highest qualification				
None	11	11	10	11
Degree level or above,	33	33	35	34
Below degree level	55	56	55	55
Urban/rural indicator				
Urban	81	82	83	83
Town	7	6	6	6

Village/Hamlet	8	9	8	8
NI	3	3	3	3
IMD - country specific quintiles				
1 - least deprived	16	16	14	15
2	17	18	18	20
3	18	19	20	20
4	24	23	24	23
5 - most deprived	25	24	24	23
Population density quintiles (from sampling frame)				
1 - least dense	23	23	22	22
2	16	16	14	16
3	20	20	18	17
4	21	21	23	24
5 - most dense	21	21	23	21
Child has a skinfold measure				
No	na	2	1	1
Yes		98	99	99
Mother weight taken				
No	na	6	6	5
Yes		95	94	95
Mother height taken				
No	na	6	6	5
Yes		94	94	95
<i>Unweighted bases</i>	<i>2212</i>	<i>973</i>	<i>513</i>	<i>611</i>

¹ Clinic-specific interview weighting factor does not include Scottish boost

Table B.7

Non-response model for DNSIYC urine sample weighting factors

	B	S.E.	Wald	df	Sig.	Exp(B)
Wave of sampling	0.20	0.140	2.1	1	0.152	1.22
In receipt of HS vouchers	0.11	0.198	0.3	1	0.568	1.12
Mothers age at interview			7.6	4	0.109	
<25 years					(baseline)	
25-29 years	0.18	0.209	0.8	1	0.381	1.20
30-34 years	0.15	0.220	0.5	1	0.489	1.16
35-39 years	0.13	0.240	0.3	1	0.588	1.14
40+ years	0.97	0.360	7.3	1	0.007	2.64
Child's age at interview			4.8	3	0.186	
4-6 months					(baseline)	
7-9 months	-0.39	0.251	2.4	1	0.124	0.68
10-12 months	0.00	0.255	0.0	1	0.994	1.00
13-18 months	-0.23	0.239	0.9	1	0.341	0.80
Child's gender			0.1	1	0.713	
Male					(baseline)	
Female	-0.05	0.140	0.1	1	0.713	0.95
Government Office Region			35.7	11	0.000	
East Midlands					(baseline)	
East of England	-0.23	0.343	0.5	1	0.497	0.79
London	0.56	0.328	2.9	1	0.088	1.75
North East	0.85	0.491	3.0	1	0.083	2.34
North West	-0.20	0.331	0.4	1	0.548	0.82
Northern Ireland	-0.53	0.470	1.3	1	0.262	0.59
Scotland	-0.56	0.358	2.4	1	0.118	0.57
South East	-0.36	0.320	1.3	1	0.259	0.70
South West	-0.26	0.356	0.5	1	0.466	0.77
Wales	-0.64	0.417	2.3	1	0.126	0.53
West Midlands	0.13	0.347	0.1	1	0.708	1.14
Yorkshire and The Humber	-0.82	0.347	5.5	1	0.019	0.44
Flag couple households			5.4	1	0.020	
Lone parent					(baseline)	
Couple	0.86	0.372	5.4	1	0.020	2.37
Number of adults aged 16 and over in household			5.4	2	0.067	
1					(baseline)	
2	0.69	0.394	3.1	1	0.081	0.50
3+	1.04	0.450	5.3	1	0.021	0.35
Mother's highest qualification			5.6	2	0.060	
None					(baseline)	
Degree level or above,	0.59	0.262	5.0	1	0.025	1.80
Below degree level	0.51	0.228	4.9	1	0.027	1.66
Constant	-0.05	0.423	0.0	1	0.907	0.95

The response is 1 = child gave a usable urine sample, 0 = child did not give a usable urine sample

Only variables that are significant at the 0.05 level are included in the model

The model R2 is 0.077 (Cox and Snell)

B is the estimate coefficient with standard error **S.E.**

The Wald-test measures the impact of the categorical variable on the model with the appropriate number of degrees of freedom df.

If the test is significant (sig. < 0.05) then the categorical variable is considered to be 'significantly associated' with the response variable and therefore included in the model