Public Health England

Protecting and improving the nation's health

Serological Surveillance: Summary report 12

Key findings:

This week's report provides:

- **Analysis of an additional :-**
    - **1) 2001 adult blood donor samples (aged 17-69 years) (collected in late June (weeks 26 and 27)) from the East of England and South East regions.**
    - **2) 641 residual sera from the SEU collected between weeks 18 and 24 (1st May -10th June) amongst adult aged 20-64 years.**
    - **3) 3004 samples collected through the Royal College of General Practitioners Research and Surveillance Centre (RCGP) participating practices amongst adults aged 20-64 years attending primary care for routine blood tests between weeks 19 and 24 (1st May and 8th June).**
    - **4) 244 paediatric residual samples from Great Ormond Street Hospital (GOSH) for weeks 16 to 21 (14 Apr – 20 May)**
    - **5) 122 paediatric residual samples from the SEU and paediatric collections collected from weeks 20 to 24 (1st May to 10th June)**
    - **6) 314 paediatric samples from the What's the Story study between weeks 8 to 22 (17th Feb and 29th May)**
- **Adjusted prevalence amongst blood donors in weeks 26-27 has plateaued in both the East of England and the South East regions, at 4.8% and 4.4% respectively.**
- **The trends in prevalence estimates from RCGP collection are consistent by region and age group with blood donor samples with both showing the highest prevalence in London and in young adults. Prevalence estimates from the SEU collection are generally higher with an increased prevalence in those aged 60-64 years compared with other adult age groups.**
- **Overall prevalence estimates in children and young adults are similar for the SEU /paediatric collections and 'What's the STORY' collection in May (approximately 5%) with considerably higher estimates from GOSH samples.**

This week's report includes additional data from testing adult samples from blood donors in England (NHS Blood and Transplant (NHSBT) with regions sampled at different time periods. We present results from testing a third set of samples from the East of England, and the third set of samples from the South East (comprising 2001 new samples in total). We also present the results from testing 641 residual samples from individuals aged 20-64 years collected through the PHE Sero Epidemiology Unit (SEU) and 3004 samples from the same age group collected from patients attending for routine blood tests at RCGP RSC participating practices across England; both sample sets were collected between 1st May and 8th June.

We also present the results of testing samples from three paediatric /young adult collections including i) 244 residual samples from individuals under 18 years from Great Ormond Street Hospital (GOSH) for weeks 14-21 ii) 122 samples from individuals aged less than 20 years from the SEU and paediatric collections (collected from 1st May to 9th June) and iii) 444 paediatric samples from the What's the Story study from participants aged under 25 years collected between 17th Feb and 29th May.

Results

Blood donor data (aged 17-69 years)

Seroprevalence estimates presented here are based on a total of 25474 adult samples from NHSBT and Welsh Blood Service (WBS) and includes the results of 989 new samples from the East of England and 1012 new samples from the South East (collected between the 26th and 29th of June (weeks 26 and 27)).

Seroprevalence estimates amongst blood donors were adjusted for the sensitivity and specificity of the EuroImmun assay, based on sensitivity of 132/160 (82.5%) and specificity of 569/574 (99.1%) and uncertainty using a Bayesian approach.

National prevalence estimates

Samples are collected from each region once every four weeks, except London where samples are collected once every two weeks. We therefore produce national prevalence estimates based on a rolling 4-weekly period.

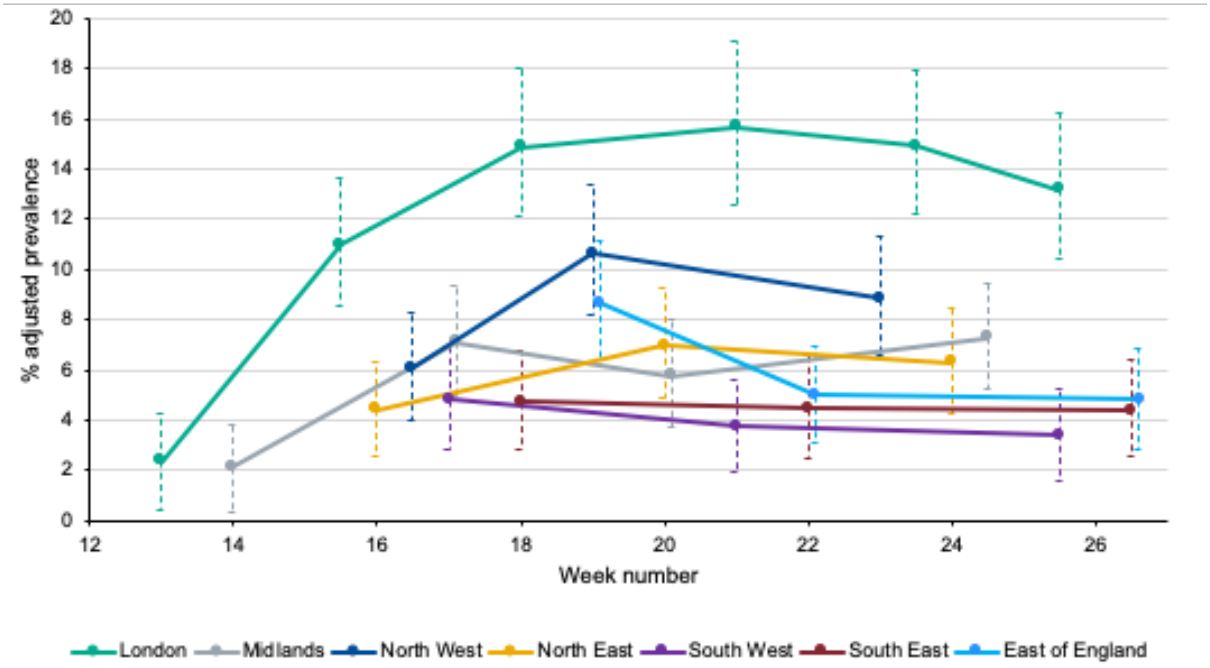
Overall population weighted (by age group, sex, NHS region) prevalence among blood donors was 7.2% (95% CI 6.6% - 7.8%) (unadjusted) or 7.6% (95% CrI 6.9% - 8.4%) after adjustment for sensitivity and specificity for the period 4 – 29 June (weeks 23-27). This compares with 7.8% (95% CI 7.2% - 8.6%) (unadjusted) or 8.3% (95% CrI 7.5% - 9.2%) (adjusted) for the period of 6 – 29 May (weeks 19-22).

Regional prevalence estimates (unweighted) over time

The additional results from weeks 26 and 27 (**Figure 1**) show that adjusted prevalence in the East of England has plateaued, at 5.0% (95% CrI 3.1% - 6.9%) in week 22 and 4.8% (95% CrI 2.9% - 6.8%) in week 27. The week 27 data for the South East (the third sample set from this

region) also indicates a plateau; adjusted prevalence has remained at 4.4% (95% CrI 2.5% - 6.4%) between weeks 22 and 27. (**Table 1, Appendix 1**). These trends are consistent with recent data from other regions.

Figure 1: *Adjusted SARS-CoV-2 antibody seroprevalence in UK blood donors*



*using Euroimmun assay adjusted for sensitivity (82.5%) and specificity (99.1%)
**error bars show 95% confidence intervals

Please note that about 1-2% of the samples come without demographic data, and hence prevalence estimates in this report are based on the 98-99% sets with available data.

Seroprevalence in adults aged 20-64 years (SEU and RCGP collections)

As a comparison with the results from testing blood donors, two additional collections of indviduals aged 20-64 have been analysed; these are SEU residual sera from participating hospital laboratories across the country and RCGP sera collected via general practioners at the time of routine blood tests. These include 641 samples from the SEU collected between 1st May and 10th June and 3004 samples from the RCGP collected between 1st May and 8th June.

These samples have been tested using the Abbott assay and adjusted for the sensitivity and specificity of the Abbott assay, based on sensitivity of 91.4% and specificity of 99.8%. The NHSBT samples were tested using the Euroimmun assay.

Overall and regional prevalence estimates from the SEU and RCGP collections are given in Tables 1a and 1b respectively. Note that the number of regional samples is sometimes small, and intervals around estimates can be wide. Table 1c additionally gives comparable NHSBT estimates based on the EuroImmun assay over the same period (1st May – 10th June).

When compared to the NHSBT and RCGP data from the same time period, prevalence estimates for the SEU collection are significantly higher in almost all regions. In comparison, the RCGP prevalence estimates are similar to the NHSBT estimates for the same time period (**Table 1c**).

***Table 1a**: Summary of SEU (unweighted) Prevalence Estimates (ages 20 – 64) using the Abbott assay, all samples were taken between 1 May and 10 June.*

region	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
All regions	62	8	571	641	9.7% (7.5% - 12.2%)	10.4% (8% - 13.2%)
EE	6	0	59	65	9.2% (3.5% - 19.0%)	10.1% (4% - 19.7%)
London	5	0	16	21	23.8% (8.2% - 47.2%)	26.5% (10.5% - 49%)
Midlands	13	0	64	77	16.9% (9.3% - 27.1%)	18.5% (10.5% - 28.9%)
NE	11	0	173	184	6.0% (3.0% - 10.4%)	6.4% (3.3% - 11%)
NW	20	0	173	193	10.4% (6.4% - 15.6%)	11.2% (7% - 16.6%)
SE	1	0	36	37	2.7% (0.1% - 14.2%)	3.1% (0.1% - 12.8%)
SW	6	0	58	64	9.4% (3.5% - 19.3%)	10.3% (4.1% - 20%)

***Table 1b**: Summary of RCGP (unweighted) Prevalence Estimates (ages 20 – 64) using the Abbott assay, all samples were taken between 1 May and 8 June.*

region	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
All regions	157	28	2819	3004	5.2% (4.5% - 6.1%)	5.5% (4.6% - 6.5%)
EE	4	0	52	56	7.1% (2.0% - 17.3%)	7.9% (2.4% - 17.4%)
London	17	0	139	156	10.9% (6.5% - 16.9%)	11.8% (7.1% - 18%)
Midlands	22	0	338	360	6.1% (3.9% - 9.1%)	6.5% (4.1% - 9.6%)
NE	20	0	377	397	5.0% (3.1% - 7.7%)	5.3% (3.2% - 8.1%)
NW	27	0	483	510	5.3% (3.5% - 7.6%)	5.6% (3.7% - 8%)
SE	20	0	386	406	4.9% (3.0% - 7.5%)	5.2% (3.2% - 7.9%)
SW	42	0	983	1025	4.1% (3.0% - 5.5%)	4.3% (3% - 5.8%)

***Table 1c**: Summary of NHSBT Prevalence Estimates by region (of residence, not sampling), using the Euroimmun assay. All samples were taken between 1 May and 10 June.*

region	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
All regions	1580	227	20621	22428	7.0% (6.7% - 7.4%)	7.5% (6.2% - 8.6%)
EE	176	25	2707	2908	6.1% (5.2% - 7.0%)	6.3% (4.7% - 7.7%)
London	562	71	4404	5037	11.2% (10.3% - 12.1%)	12.6% (10.9% - 14.3%)
Midlands	210	31	3311	3552	5.9% (5.2% - 6.7%)	6.1% (4.6% - 7.5%)
NE	162	22	2561	2745	5.9% (5.0% - 6.8%)	6.1% (4.5% - 7.6%)
NW	216	29	2395	2640	8.2% (7.2% - 9.3%)	8.9% (7.2% - 10.6%)
SE	139	27	2523	2689	5.2% (4.4% - 6.1%)	5.2% (3.7% - 6.6%)
SW	115	22	2720	2857	4.0% (3.3% - 4.8%)	3.8% (2.3% - 5%)

Seroprevalence estimates in individuals under the age of 25

The results of testing three different collections with samples collected in children and young adults are presented. This includes 244 residual sera from Great Ormond Street Hospital (GOSH) for weeks 14-21, 122 additional residual samples from the SEU and paediatric hospital collections collected from 1st May to 9th June and 444 samples from the What's the Story collection between 17th Feb and 29th May **(Table 2)**. What's the Story is a representative household survey that has been collecting sera from healthy children and adolescents under the age of 25 years in England. Further details have been provided in previous reports.

A subset of SEU and paediatric samples were tested using a PHE in house Receptor Binding Domain (RBD) assay – this was limited to the samples that had sufficient available volume once testing wih the Abbott assay was completed. Seroprevalence estimates from the RBD assay were adjusted for a sensitivity of 149/161 (92.9%) and specificity of 1100/1122 (98.0%)

***Table 2a**: Summary of GOSH (unweighted) Prevalence Estimates (ages 1 – 18) by period of sampling, using the Euroimmun assay*

date range	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
20-28 Mar	14	7	169	190	7.4% (4.1% - 12.1%)	7.9% (3.9% - 13.4%)
1-30 Apr	105	18	601	724	14.5% (12.0% - 17.3%)	16.7% (13.4% - 20.4%)
1-20 May	72	10	432	514	14.0% (11.1% - 17.3%)	16.1% (12.4% - 20.3%)

***Table 2b**: Summary of SEU and Paediatric (unweighted) Prevalence Estimates (ages 1 – 19) by period of sampling, using the Abbott and RBD assays*

date range	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
Abbott assay						
1 Feb - 31 Mar	2	6	435	443	0.5% (0.1% - 1.6%)	0.3% (0% - 1.3%)
1 - 30 Apr	39	4	604	647	6.0% (4.3% - 8.1%)	6.4% (4.5% - 8.7%)
1 May - 9 Jun	34	9	675	718	4.7% (3.3% - 6.6%)	5% (3.4% - 6.9%)
RBD assay						
1 Feb - 31 Mar	5	3	305	313	1.6% (0.5% - 3.7%)	0.2% (0% - 1.9%)
1 - 30 Apr	37	7	467	511	7.2% (5.1% - 9.8%)	5.8% (3.3% - 8.6%)
1 May - 9 Jun	27	5	484	516	5.2% (3.5% - 7.5%)	3.5% (1.3% - 6.1%)

***Table 2c**: Summary of What's the Story (unweighted) Prevalence Estimates (ages 1 – 24) by period of sampling, using the Abbott assay*

date range	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
17 Feb - 31 Mar	0	1	107	108	0.0% (0.0% - 3.4%)	0.2% (0% - 2.5%)
1 - 30 Apr	9	2	201	212	4.2% (2.0% - 7.9%)	4.5% (2.1% - 8.1%)
1 - 29 May	5	0	119	124	4.0% (1.3% - 9.2%)	4.3% (1.4% - 9.3%)

Prevalence estimates for GOSH are consistently higher than other two collections – this is likely explained by location of sampling; patients at GOSH live predominantly in London and the neighbouring regions where prevalence is higher. In comparison, the What's the Story and SEU collections have comparable prevalence estimates.

All three sample sets show a similar pattern of prevalence among paediatric patients – an increase in prevalence between February and early April, followed by a plateau or slightly lower prevalence between April and late May.

National Seroprevalence estimates, by age group (NHSBT, SEU and RCGP collections)

We present the analysis of all available results tested using the Abbott assay from the SEU (and Paediatric) and RCGP datasets that are weighted by age group and region.

SEU and Paediatric collection samples were taken between 1st May and 10th of June and includes individuals aged 1 to 64. RCGP samples were taken between 1st May and 8th June and includes individuals aged 11 to 110, analyses of RCGP samples in individuals aged 65+ were presented in a previous report. All estimates, both unadjusted and adjusted, are

derived via Bayesian multilevel regression models (with the aim of producing stable estimates for each age group and region combination even where data are sparse) with poststratification.

National prevalence for the SEU and Paediatric collections was 8.3% (6.5% - 10.6%) (unadjusted) or 8.9% (6.9% - 11.4%) (adjusted for sensitivity & specificity of the Abbott assay). National prevalence for the RCGP collections was 5.7% (4.6% - 7%) (unadjusted) or 6% (4.8% - 7.5%) (adjusted for sensivity & specificity of the Abbott assay). This compares with estimates from testing blood donors using the Euroimmun assay over the same period (1st May – 10th June) at 7.6% (7% - 8.1%) (unadjusted) or 8.1% (7.4% - 8.8%) (adjusted for sensitivity & specificity of the EuroImmun assay).

When stratified by age, the RCGP point prevalence estimates display a similar pattern to the results from testing blood donor samples **(Figure 2c)**; young adults in particular display a significantly higher prevalence than older age bands (**Figure 2a**). A different pattern however is observed from testing the combined SEU and paediatric collections with the highest prevalence observed among people over the age of 60. (**Figure 2b).**

***Figure 2a**: Region weighted % pos (with 95% CI) by age group, RCGP collection data, 1st May to 8th June*

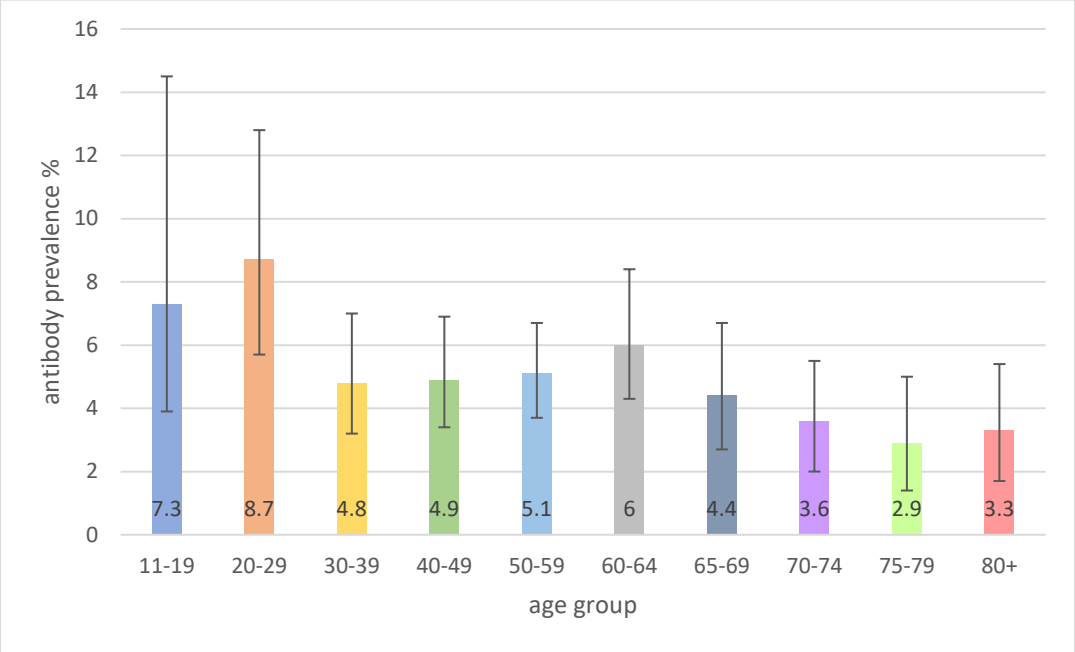


***Figure 2b**: Region weighted % pos (with 95% CI) by age group, SEU and Paediatric collection data, 1st May – 10th June*

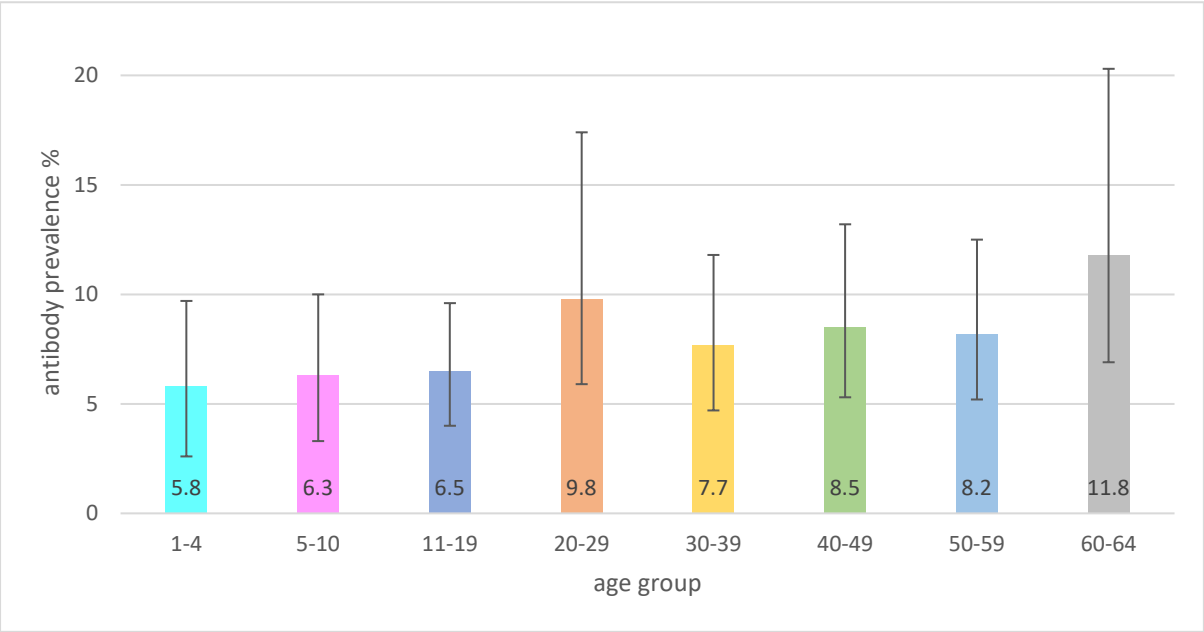
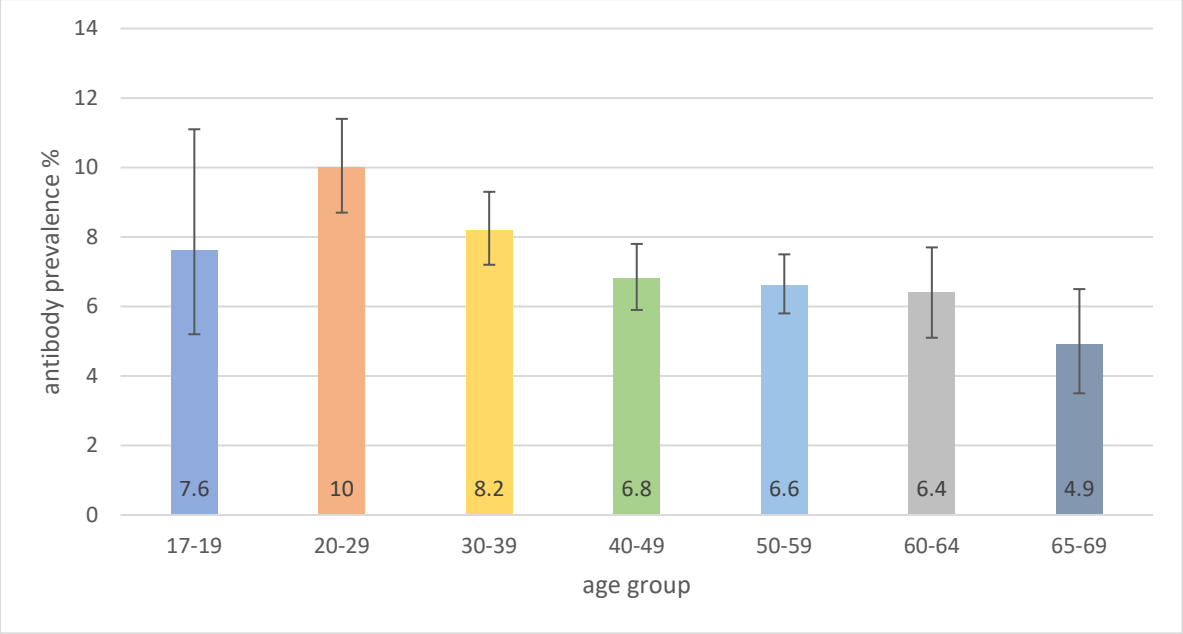


***Figure 2c**: Region weighted EuroImmun % positive (with 95% CI) by age group in blood donors, 1st May – 10th June*



Comments

We report national prevalence estimates using three different collections for the period of early May to early June. We found that the weighted unadjusted prevalence was 7.6% (95% CrI 6.9% - 8.4%) based on testing blood donor samples; 5.7% (4.6% - 7%) from RCGP collection and 8.9% (6.9% - 11.4%) based on testing residual sera from the SEU and Paediatric collections.

Data from the RCGP and SEU collections are also stratified by age group. In line with the analyses of NHSBT data, prevalence in the RCGP collections indicated a decreasing prevalence with increasing age. A somewhat different pattern is observed from the testing of residual sera from the SEU/paediatric collection, with the highest prevalence in the older adults. This likely reflects the different sources of the samples – the SEU dataset is collected from patients in hospital, a group in which older people who are SARS-CoV-2 positive are likely to be overrepresented.

Updated regional prevalence estimates based on the results of testing the third set of donor samples from the East of England and a third in the South East (weeks 26-27) are presented. The latest results are consistent with prevalence reaching a plateau in both regions at 4.8% and 4.4% respectively. These trends are generally consistent with the pattern observed in other regions although some regions have observed lower prevalence in more recent periods. However, any changes in prevalence estimates over time also need to take account of potential changes in the precise locations of sampling in each region over time, assay variability and waning antibody. From the beginning of June individuals have been advised not to donate blood for 28 days following resolution of symptoms suggestive of COVID-19, an increase from previous exclusion of 14 days.

Prevalence amongst children appear to be generally lower than young adults with similar estimates derived from testing samples from the SEU /Paediatric collections and Whats the STORY study, at around 5% during May. In comparion, prevalence estimates from GOSH are considerably higher. This is likely explained – at least in part - by location of sampling; patients at GOSH live predominantly in London and the neighbouring regions where prevalence is higher than in the rest of the population. In addition, the GOSH sample set comprises a specialist patient population which have higher rates of health care attendance (and therefore potentially exposure) than general paediatric population. Despite these differences, trends across the three collections are similar with prevalence reaching a plateau, consistent with the trends observed in adults.

Appendix 1: NHSBT data

***Table 1**: Summary of NHSBT Prevalence Estimates by region and period of sampling, using the Euroimmun assay*

Region	date range	week	pos	ind	neg	total	% pos (95% CI)	adjusted prevalence (95% CrI)
NHS blood & transplant								
London	26-27 Mar	13	22	11	724	757	2.9% (1.8% - 4.4%)	2.3% (0.4% - 4.3%)
	9-13 Apr	15-16	107	15	963	1085	9.9% (8.2% - 11.8%)	11% (8.6% - 13.6%)
	1-3 May	18	127	10	837	974	13.0% (11.0% - 15.3%)	14.9% (12.1% - 18%)
	21-22 May	21	109	21	667	797	13.7% (11.4% - 16.3%)	15.6% (12.6% - 19.1%)
	5-8 Jun	23-24	143	9	942	1094	13.1% (11.1% - 15.2%)	14.9% (12.2% - 17.9%)
	19-22 Jun	25-26	106	12	793	911	11.6% (9.6% - 13.9%)	13.1% (10.4% - 16.2%)
Midlands	2-3 Apr	14	25	13	878	916	2.7% (1.8% - 4.0%)	2.1% (0.3% - 3.8%)
	23-24 Apr	17	70	9	964	1043	6.7% (5.3% - 8.4%)	7.1% (5% - 9.3%)
	14-15 May	20	49	6	815	870	5.6% (4.2% - 7.4%)	5.8% (3.7% - 8%)
	11 Jun	24-25	77	7	1040	1124	6.9% (5.4% - 8.5%)	7.3% (5.2% - 9.5%)
NE	14-16 Apr	16	46	12	959	1017	4.5% (3.3% - 6.0%)	4.4% (2.5% - 6.3%)
	13-14 May	20	67	8	939	1014	6.6% (5.2% - 8.3%)	7% (4.9% - 9.2%)
	10-12 Jun	24	64	5	987	1056	6.1% (4.7% - 7.7%)	6.3% (4.3% - 8.4%)
NW	15-20 Apr	16-17	55	11	870	936	5.9% (4.5% - 7.6%)	6.1% (4% - 8.3%)
	6-8 May	19	92	16	852	960	9.6% (7.8% - 11.6%)	10.6% (8.2% - 13.3%)
	4-6 Jun	23	83	6	931	1020	8.1% (6.5% - 10.0%)	8.8% (6.6% - 11.3%)
SW	24-26 Apr	17	42	8	815	865	4.9% (3.5% - 6.5%)	4.8% (2.8% - 6.9%)
	21-22 May	21	42	14	994	1050	4.0% (2.9% - 5.4%)	3.7% (1.9% - 5.6%)
	19 – 22 Jun	25-26	38	4	970	1012	3.8% (2.7% - 5.1%)	3.4% (1.6% - 5.2%)
SE	30 Apr - 1 May	18	49	11	960	1020	4.8% (3.6% - 6.3%)	4.7% (2.8% - 6.7%)
	28-29 May	22	38	7	787	832	4.6% (3.3% - 6.2%)	4.4% (2.5% - 6.6%)
	26-29 Jun	26-27	46	7	959	1012	4.5% (3.3% - 6.0%)	4.4% (2.5% - 6.4%)
EE	7-10 May	19	81	13	921	1015	8.0% (6.4% - 9.8%)	8.7% (6.4% - 11.1%)
	28-29 May	22	55	6	1039	1100	5.0% (3.8% - 6.5%)	5% (3.1% - 6.9%)
	26-29 Jun	26-27	48	3	938	989	4.9% (3.6% - 6.4%)	4.8% (2.9% - 6.8%)
Welsh blood service								
Wales		17	34	4	968	1006	3.4% (2.4% - 4.7%)	3% (1.1% - 4.7%)

Appendix 2: RCGP and SEU data

***Table 2a**: Summary of SEU collection age group and region weighted prevalence Estimates overall and by age, using the Abbott assay, date range 1 May – 10 June.*

group	pos	total	Modelled population weighted % pos (95% CrI)	Modelled population weighted adjusted prevalence (95% CrI)
overall	93	1284	8.3% (6.5% - 10.6%)	8.9% (6.9% - 11.4%)
age 1-4	5	137	5.8% (2.6% - 9.7%)	6.1% (2.6% - 10.3%)
age 5-10	8	172	6.3% (3.3% - 10%)	6.7% (3.3% - 10.8%)
age 11-19	18	342	6.5% (4% - 9.6%)	6.9% (4.1% - 10.4%)
age 20-29	9	72	9.8% (5.9% - 17.4%)	10.5% (6.2% - 18.9%)
age 30-39	12	163	7.7% (4.7% - 11.8%)	8.2% (4.9% - 12.8%)
age 40-49	13	145	8.5% (5.3% - 13.2%)	9.1% (5.6% - 14.3%)
age 50-59	14	168	8.2% (5.2% - 12.5%)	8.8% (5.5% - 13.5%)
age 60-64	14	85	11.8% (6.9% - 20.3%)	12.8% (7.4% - 22.1%)

***Table 2b**: Summary of RCGP collection age group and and region weighted prevalence Estimates overall and by age, using the Abbott assay, date range 1 May – 8 June.*

group	pos	total	Modelled population weighted % pos (95% CrI)	Modelled population weighted adjusted prevalence (95% CrI)
overall	198	4387	5.7% (4.6% - 7%)	6% (4.8% - 7.5%)
age 11-19	7	68	7.3% (3.9% - 14.5%)	7.8% (4% - 15.9%)
age 20-29	27	297	8.7% (5.7% - 12.8%)	9.4% (6% - 13.9%)
age 30-39	21	488	4.8% (3.2% - 7%)	5.1% (3.2% - 7.5%)
age 40-49	28	631	4.9% (3.4% - 6.9%)	5.2% (3.5% - 7.4%)
age 50-59	44	946	5.1% (3.7% - 6.7%)	5.3% (3.8% - 7.2%)
age 60-64	32	548	6% (4.3% - 8.4%)	6.4% (4.4% - 9%)
age 65-69	14	354	4.4% (2.7% - 6.7%)	4.6% (2.7% - 7.1%)
age 70-74	11	389	3.6% (2% - 5.5%)	3.6% (2% - 5.9%)
age 75-79	6	326	2.9% (1.4% - 5%)	3% (1.3% - 5.2%)
age 80+	8	340	3.3% (1.7% - 5.4%)	3.4% (1.7% - 5.6%)

***Table 2c**: Summary of NHSBT collection age group and region weighted prevalence Estimates overall and by age, using the EuroImmun assay, date range 1 May – 10 June.*

group	pos	total	Modelled population weighted % pos (95% CrI)	Modelled population weighted adjusted prevalence (95% CrI)
overall	905	11228	7.6% (7% - 8.1%)	8.1% (7.4% - 8.8%)
age 17-19	14	157	7.6% (5.2% - 11.1%)	8.1% (5.2% - 12.5%)
age 20-29	221	1854	10% (8.7% - 11.4%)	11.1% (9.4% - 12.9%)
age 30-39	209	2329	8.2% (7.2% - 9.3%)	8.8% (7.5% - 10.3%)
age 40-49	163	2291	6.8% (5.9% - 7.8%)	7.2% (6% - 8.5%)
age 50-59	191	2751	6.6% (5.8% - 7.5%)	7% (5.9% - 8.2%)
age 60-64	74	1113	6.4% (5.1% - 7.7%)	6.7% (5.1% - 8.4%)
age 65-69	33	733	4.9% (3.5% - 6.5%)	4.9% (3.2% - 6.9%)

Appendix 3: statistical methods

The unweighted observed prevalence, $prev_{obs}$, is calculated as n^+/N, where n^+ is the number of individuals who tested positive and N is the total number of individuals tested with an available result. 95% exact confidence intervals were calculated for $prev_{obs}$ in STATA (version 14).

Population weighted observed prevalences for NHSBT data were calculated using svy commands with the poststrata() option in STATA (version 14). See the STATA Survey Data Refence Manual for methodological details.

It is understood that all assays are imperfect and can sometimes give false positive and false negative results, with probability (1-*Sp*) and (1-*Se*) respectively, where *Sp* denotes the Specificity or the probability that the test gives a negative result in individuals who have not experienced the disease, and *Se* denotes the Sensitivity or the probability that the test gives a positive result in individuals who have experienced the disease. The adjusted prevalence, denoted $prev_{adj}$, should better reflect the proportion of the population that have experienced the disease; this is related to the observed prevalence as follows:

$$prev_{obs} = Se \times prev_{adj} + (1 - Sp) \times \left(1 - prev_{adj}\right)$$

(see Diggle 2011, Lewis & Torgerson 2012). This relation was incorporated in a Bayesian model, along with the sampling distribution for positive tests $n^+ \sim \text{Binomial}(N, prev_{obs})$. The sensitivity and specificity are not known exactly, but are informed by data. Counts of true positives and false negatives in convalescent sera were used to estimate the sensitivity, and similarly counts of true negatives and false positives in pre-covid19 baseline sera were used to estimate the specificity. The sensitivity, *Se*, and specificity, *Sp*, were included in our Bayesian model each by way of a conjugate Beta-Binomial model with a Beta(0.5,0.5) reference prior, thus uncertainty in their true value was taken into account.

In unweighted adjustment models, we use a Beta(0.5,0.5) (Jeffreys) prior for the adjusted prevalence $prev_{adj}$. MCMC models were run using the NIMBLE package in R, default sampler, 500,000 iterations with a burn-in of 1,000 iterations and a thinning interval of 5.

Models to estimate population weighted $prev_{adj}$, were further extended to a multilevel logistic regression model, including a random effect for age and region specific seroprevalences (plus a fixed effect for gender when modelling the NHSBT data), following Park et al (2004)'s multilevel regression and poststratification (MRP) models. If each `cell' combination of age and region (and gender, if included) is denoted *j*, then the weighted or poststratified prevalence is given by

$$prev_{weighted} = \frac{\sum_j N_j prev\_adj_j}{\sum_j N_j}$$

Where N_j denotes the population of each cell taken from ONS data. Similar models were used to estimate population weighted observed prevalence for the SEU and RCGP data. MCMC models were run using STAN and the rstan package in R, 4 chains of length 25,000, with a burn-in of 1,000 iterations.

References

Diggle PJ. Estimating prevalence using an imperfect test. Epidemiology Research International, 2011, ArticleID 608719.

Lewis FI and Torgerson PR. A tutorial in estimating the prevalence of disease in humans and animals in the abcence of a gold standard diagnostic. Emerg Themes Epidemiol, 2012, 9: 9.

StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.

STATA Survey Data Reference Manual. Stata Press, 2019. ISBN-13: 978-1-59718-298-1. https://www.stata.com/bookstore/survey-data-reference-manual/ (pages 62-65)

NIMBLE Development Team. 2020. NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling. doi: 10.5281/zenodo.1211190. R package version 0.9.1, https://cran.r-project.org/package=nimble.

Park DK, Gelman A and Bafumi J. Bauesian multilevel estimation with poststratification: state level estimates from national polls. Political analysis, 2004, 12:375-385.

Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. Journal of Statistical Software 76(1). DOI 10.18637/jss.v076.i01

Stan Development Team. 2018. RStan: the R interface to Stan. R package version 2.17.3. http://mc-stan.org

Office for National Statistics. 2019. Population estimates by output areas, electoral, health and other geographies, England and Wales: mid-2018. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualsmallareapopulationestimates/mid2018