User Guide: Pgp3 antibodies in Health Survey for England participants
(Aggregate dataset)

We request that users contact the study coordinator, Dr Sarah Woodhall
(sarah.woodhall@phe.gov.uk), or Chief Investigator, Dr Anthony Nardone
(anthony.nardone@phe.gov.uk) prior to use of these data to ensure correct
interpretation.

Overview of the dataset

Number of observations: 5618
Data prepared by: Sarah Woodhall (sarah.woodhall@phe.gov.uk)
Filename(s): HSE pgp3 HSV2 antibodies_for UK Data Archive.dta ; HSE pgp3 HSV2
antibodies_for UK Data Archive.csv

This dataset includes a subset of data from the Health Survey for England (HSE)
(http://www.hscic.gov.uk/healthsurveyengland), which has been combined with results of
testing of participants’ blood samples for antibodies to Chlamydia trachomatis (chlamydia).

The dataset has been generated as a result of an NHS ethics committee-approved study
entitled ‘Estimation of the seroprevalence of antibodies to Chlamydia trachomatis and
Herpes Simplex Virus type 2 in the English population’ (REC ref: 13/YH/0304; Chief
Investigator: Dr Anthony Nardone, Public Health England). Results from this study are
presented in Woodhall et al Chlamydia trachomatis Pgp3 antibody population
seroprevalence before and during an era of widespread opportunistic chlamydia screening in

Consent and anonymization:
Since 1994, blood samples have been collected, within the nurse visit, as part of the HSE. At
the time of taking the blood sample, participants provided their written consent for their
stored blood samples to be tested as part of other, Research Ethics Committee-approved
studies. Only samples from participants who gave written permission, both for storage and
future analyses of a blood sample, are included in the study. Samples provided before 2002
did not provide consent for samples to be tested for viruses. From 2002 onwards, consent was provided for storage and further analysis (excluding HIV virus testing). Therefore only samples from participants who provided samples from 2002 onwards had their samples tested for HSV2.

All sera were tested anonymously using unique study identifiers, such that test results could not be linked back to any personal identifiable information.

The identifier used in this dataset is unique to this study and is unlinked from the main HSE dataset provided on the UK Data Archive. As such no attempt should be made to link this dataset to the main HSE dataset.

**Population:**
Data are from HSE participants who had provided a blood specimen with informed consent for future use and include data from two groups:
- a) 16 to 44 year-old male and female participants in HSE2010 and HSE2012
- b) Female HSE participants aged 16 to 24 years who took part in HSE in years when stored sera were available (1994-1996, 2001-02 and each year from 2008-2012).

A subset of data from HSE questionnaire items is included in the dataset to allow investigation in relation to demographics and sexual behaviours. Sexual behaviour data were only collected in the HSE in 2010 and 2012.

Queries about the dataset or its use should be directed to Sarah Woodhall, Public Health England (sarah.woodhall@phe.gov.uk).
Details of biological testing

*C. trachomatis* antibody testing

Stored sera were tested for *C. trachomatis* antibodies using two Pgp3 ELISAs. The testing strategy used two ‘in-house’ immunoglobulin G (IgG) enzyme-linked immunosorbent assays (ELISAs) based on the *C. trachomatis*-specific antigen Pgp3 (testing performed at Imperial College London). All specimens were tested using an indirect ELISA, the performance characteristics of which have been previously described. Briefly, sensitivity to detect a previous known infection was assessed among women and men with a clinical diagnosis of chlamydia at least one month previously and was found to be 73.8% (66.5–79.9) in women and 44.2% (37.3–51.3) in men. Specificity was estimated using microimmunfluourescence (MIF)-negative paediatric sera to reduce likelihood of a previous sexually-acquired *C. trachomatis* infection, and was found to be 97.6% (95%CI 96.2–98.6%). The second assay used in our testing strategy was a double-antigen sandwich ELISA (hereafter ‘double-antigen ELISA’), which has demonstrated equivalent specificity (97.8%, 95% CI 96.5–99.1), but higher sensitivity (82.9%, 77.0–88.8 in women; 54.4%, 47.2–61.6 in men) than the Pgp3 indirect ELISA against the same clinical samples. Although the double-antigen ELISA has demonstrated higher sensitivity than the indirect ELISA, the double-antigen ELISA requires around a 25-fold higher volume of sera. A separate comparison of results from sera tested on both assays showed that the indirect ELISA has good agreement with the double-antigen ELISA at low (<0.1) and high (>1.0) absorbance values. The indirect ELISA was therefore used for initial screening, with subsequent testing of sera with absorbance values between 0.1 and 1.0 using the double-antigen ELISA to resolve ‘equivocal’ specimens (Table 1).

Results of both assays as well as the combined final result are provided in the dataset.

### Table 1: Pgp3 antibody test result according to testing strategy

<table>
<thead>
<tr>
<th>Indirect ELISA result (Absorbance range)</th>
<th>Double-antigen sandwich ELISA result</th>
<th>Double-antigen sandwich ELISA result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative (&lt;0.1)</td>
<td>Pgp3 negative</td>
<td>N/A: Not re-tested</td>
</tr>
<tr>
<td>Negative (0.1-0.4731)*</td>
<td>Pgp3 negative</td>
<td>Pgp3 positive</td>
</tr>
<tr>
<td>Positive (0.4731-1.0)</td>
<td>Pgp3 positive</td>
<td>Pgp3 positive</td>
</tr>
<tr>
<td>Positive (&gt;1.0)</td>
<td>Pgp3 positive</td>
<td>N/A: Not re-tested</td>
</tr>
</tbody>
</table>

*An absorbance (450-620nm) value of 0.473 is the cutoff for the indirect assay."
Weighting, clustering and stratification

Weighted and unweighted numerators and denominators are provided in the dataset. Weights reflect the weighting provided in each HSE year. Details can be found in the HSE user guides. In summary, no weights were applied in HSE before 2002. In 2002 a weight was applied in the main HSE dataset to reflect unequal probability of selection by age group. For 2008 onwards, a blood weight was generated in the main HSE dataset for all adults who had a nurse visit, were eligible for, agreed and were able to give a blood sample.

Although the data came from a complex survey sample, the aggregate nature of the data mean that no clustering or stratification has been applied.

List of variables

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable description</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>year of HSE survey</td>
</tr>
<tr>
<td>age</td>
<td>age in years</td>
</tr>
<tr>
<td>sex</td>
<td>gender</td>
</tr>
<tr>
<td>ind_pos_unwt</td>
<td>Number positive on Indirect ELISA (unweighted)</td>
</tr>
<tr>
<td>pgp3_pos_unwt</td>
<td>Number positive on Double-antigen ELISA (unweighted)</td>
</tr>
<tr>
<td>denom_unwt</td>
<td>Total number tested (unweighted)</td>
</tr>
<tr>
<td>ind_pos_wt</td>
<td>Number positive on Indirect ELISA (weighted)</td>
</tr>
<tr>
<td>pgp3_pos_wt</td>
<td>Number positive on Double-antigen ELISA (weighted)</td>
</tr>
<tr>
<td>denom_wt</td>
<td>Total number tested (weighted)</td>
</tr>
</tbody>
</table>


References