**GP Patient Survey**
**Weighting Strategy for Year 2011-2012**

**Important note – 2011-2012 Survey**

Please note that due to the changes to the questionnaire design and survey frequency, as well as the change to the weighting methodology, no results from the 2011-12 survey can be compared to previous years, even where questions remain the same.

**The Weighting Strategy for 2011-2012**

The new weighting scheme for the 2011-2012 year of the survey incorporates a weight to account for the unequal probability of selection by practice with a weight to account for differential response patterns. This new weight has been shown to provide survey estimates that are less biased than the old weighting scheme[1].

The new weight can be broken down into three parts:

1. a **design weight** to account for the unequal probability of selection;
2. a **non-response weight** to account for differences in the characteristics of responders and non-responders; and,
3. a **post-stratification** weight by practice to ensure that the weighted responding sample within each practice resembles the population of eligible patients within the practice.

**Design weights** were computed to account for the design of the survey (e.g. disproportionate stratified random sample by practice). Design weights were calculated for each practice as the inverse of the probability of selection. The probability of selection was calculated by dividing the number of selected patients over the total number of eligible patients in the practice (*excluding those patients who had been issued a questionnaire in the last 6 months*). This weight corresponds to the number of patients in a practice that is represented by each patient in the sample of that practice.

**Non-response weights** were constructed using a model based approach to estimate propensity scores. This model estimates the probability of responding based on socio-economic and demographic characteristics of the patient and the neighbourhood the patient lives in. This strategy aims to reduce demographic and socio-economic differences between respondents and non-respondents.

[1] Micklewright, John and Schnepf, Sylke V. and Skinner, Chris (2010) *Non-Response Biases in Surveys of School Children: The Case of the English PISA Samples.* Discussion Paper. Institute for the Study of Labor.

Data from the GPPS sampling frame (patient's age, gender and GOR) was linked to external data using the postcode of the patients[2]. External data was obtained from the Office of National Statistics (ONS) aggregated at the Output Area (OA), and the Classification of Residential Neighbourhoods (ACORN) system. Output Area (OA) variables included: deprivation, crime scores, ethnicity, marital status, overcrowding, household tenure and employment status. The OA variables are based on the 2001 Census data except for the Index of Multiple Deprivation (IMD) which is based on the 2010 mid-year estimates. The Classification of Residential Neighbourhoods (ACORN) system[3] categorizes all postcodes in UK into various types based on Census data and lifestyle surveys.

Some of the selected patients did not have a valid postcode or their postcode was missing. These patients were assigned the practice modal OA (the OA that most of the patients had within the practice). If there was more than one modal OA, the OA was selected randomly among the modal OAs. There was complete information for most of the patients except in ACORN group and IMD score/crime. Only one patient had a missing ACORN, this patient was assigned the practice modal Acorn Type (the Acorn Type with the most patients assigned within each practice). There were 387 patients with missing (English) IMD score/crime. The postcodes of these patients were located in Wales and they were distributed in 31 practices. In these practices, on average, 9% of the selected patients had a missing IMD. The maximum percentage of patients within a practice that had a missing IMD was 38%. The IMD score/crime for these patients was imputed using the average IMD for the practice that they attended.

The probability of response *p* was estimated using a logistic regression model. Standardised design weights were applied when running the model to obtain unbiased estimates for the coefficients. The model showed that non-response was higher among younger people and males. Furthermore, there was a significant interaction between age and gender that showed that younger males were less likely to respond than younger females; but after age 70, males were slightly more likely to respond than females. Response was lower in the North West and West Midlands than in London and it also decreased in OAs with higher deprivation and crime scores; with an increasing proportion of non-white people; with an increasing proportion of single, separated or divorced people; with an increasing proportion of households with three or more people; and with an increasing proportion of privately

[2] These variables were selected because they were significant in a previous study in GPPS Y5 (not published)
[3] http://www.caci.co.uk/acorn-classification.aspx

rented households. In contrast, response increased with an increasing proportion of employees[4].

The non-response pre-weights *pw* were calculated as the reciprocal of the predicted probability of response *p*. The pre-weights were capped at 10 to limit the decline in the precision of the survey estimates. Capping will introduce some bias into the survey estimates, however in this case it will be minimal given the number of respondents with capped weights represented 0.6% of the total respondent sample. The pre-weights were multiplied by the design weight to obtain the non-response weight *nw*.

A **post-stratification** factor was computed to make the weighted sample of respondents resemble the eligible population by practice. The practice level post-stratification factor is simply calculated as the proportion of the population in the practice divided by the proportion of responding patients (weighted by *nw*) in the practice. The non-response weight was multiplied by the post-stratification factor and standardized to obtain the final weight.

**Summary of the weights**

At the national level, the distribution of the respondents by age and gender using the final weight (figure 2 overleaf) was very similar to the distribution in the population (figure 1 overleaf). For PCTs, the average difference between the age and gender population profile and the final weighted sample varied by between 0.1% and 0.6% with an average of 0.3%.

The design effect at the national level was 1.85. At the PCT level, the design effects varied between 1.44 and 5.16 with an average of 1.79. There were only two PCTs with extremely high design effects. In one PCT there were 19 patients with weights over 10; and 14 of these patients came from one practice, a university practice. In another PCT there were 11 patients with weights over 10; and 10 of these patients came from only one practice as well, again another university practice. The large weights at the university level were due to the large design weights (low probability of selection) and low response rates.
At the practice level, the design effects varied between 1 and 1.77 with an average of 1.28.

[4] The magnitude and direction of the coefficients for this model were similar to those of the model fitted for GPPS Y5 in a previous study (not published).

**Figure 1. Eligible patient's age and gender distribution (national)**

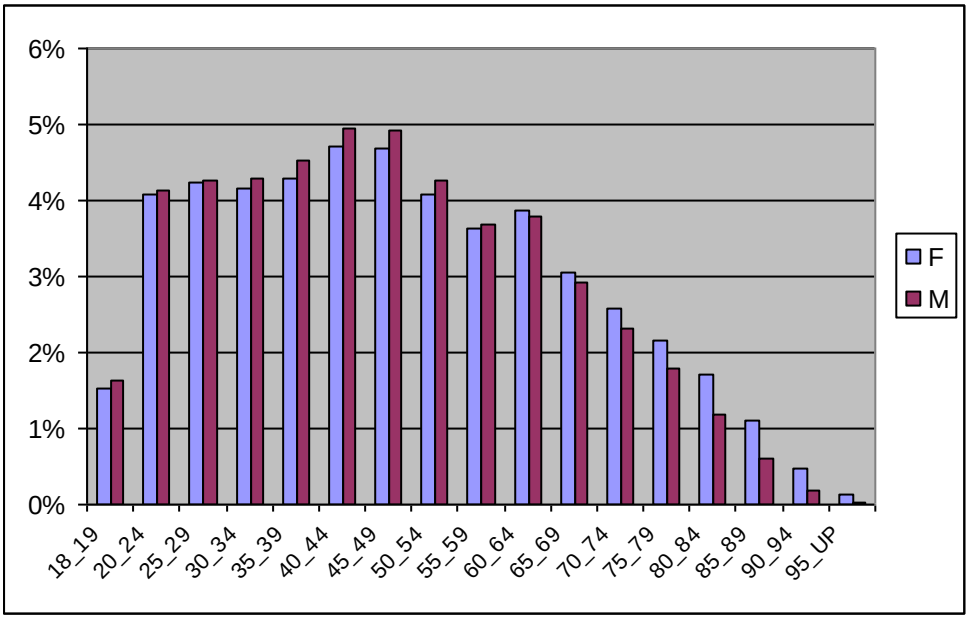


**Figure 2. Respondent's age and gender distribution (national, using final weight)**