**Health Survey for England 2014 depression analysis and local prevalence model**

**Project for Public Health England**

**Saba Mian, Miguel Pugliese, Bowen Su, Roger Newson, Bowen Su, Peter Tyrer and Michael Soljak**

**Department Primary Care & Public Health**

**School of Public Health**



**Contents**

[1 Background 4](#_Toc456104519)

[1.1 Diagnosis of depression 5](#_Toc456104520)

[1.2 Types of depression 6](#_Toc456104521)

[1.2.1 Major depressive disorder 6](#_Toc456104522)

[1.2.2 Bipolar disorder 6](#_Toc456104523)

[1.3 Risk factors 8](#_Toc456104524)

[1.3.1 Sex 9](#_Toc456104525)

[1.3.2 Age 9](#_Toc456104526)

[1.3.3 Ethnic group 9](#_Toc456104527)

[1.3.4 Body mass index (BMI) 9](#_Toc456104528)

[1.3.5 Education 9](#_Toc456104529)

[1.3.6 Marital status 9](#_Toc456104530)

[1.3.7 Socioeconomic status 9](#_Toc456104531)

[1.3.8 Anxiety 10](#_Toc456104532)

[1.3.9 Long-lasting illness 10](#_Toc456104533)

[1.3.10 Physical activity 10](#_Toc456104534)

[1.3.11 Alcohol abuse 10](#_Toc456104535)

[1.3.12 Sleep disorders 10](#_Toc456104536)

[2 Methods 13](#_Toc456104537)

[2.1 Data source & sampling 13](#_Toc456104538)

[2.1.1 Outcome variables 13](#_Toc456104539)

[2.1.2 Risk factors 13](#_Toc456104540)

[2.1.3 Missing values 14](#_Toc456104541)

[2.2 Regression modelling 14](#_Toc456104542)

[2.3 Internal validation 15](#_Toc456104543)

[2.4 Local prevalence estimates 15](#_Toc456104544)

[2.4.1 Data sources 15](#_Toc456104545)

[2.4.2 Small population estimation methods 16](#_Toc456104546)

[2.4.3 Local prevalence estimates: bootstrap method for local prevalence estimates 17](#_Toc456104547)

[2.4.4 Local prevalence estimates: inverse probability weights method 18](#_Toc456104548)

[2.5 Validation of local estimates 19](#_Toc456104549)

[2.5.1 Internal validation of local estimates 19](#_Toc456104550)

[2.5.2 External validation/comparison with QOF prevalence 19](#_Toc456104551)

[3 Results 21](#_Toc456104552)

[3.1 Population and baseline characteristics 21](#_Toc456104553)

[3.2 Statistical analysis 21](#_Toc456104554)

[3.2.1 Descriptive Statistics for the final population 21](#_Toc456104555)

[3.2.2 Missing values 23](#_Toc456104556)

[3.2.3 Univariable analyses 24](#_Toc456104557)

[3.2.4 Multivariable analyses 25](#_Toc456104558)

[3.2.5 Tests for interactions and collinearity 26](#_Toc456104559)

[3.2.6 Stepwise forward and backward selection model 27](#_Toc456104560)

[3.3 Internal validation/discrimination 27](#_Toc456104561)

[3.3.1 ROC curves 27](#_Toc456104562)

[3.3.2 Predicted probabilities of having depression 30](#_Toc456104563)

[3.3.3 Sensitivity and specificity analysis 30](#_Toc456104564)

[3.3.4 Collinearity 30](#_Toc456104565)

[3.4 Local prevalence estimates 30](#_Toc456104566)

[3.4.1 External validation of local estimates 31](#_Toc456104567)

[4 Discussion 34](#_Toc456104568)

[4.1 Summary 34](#_Toc456104569)

[4.2 Strengths and limitations 34](#_Toc456104570)

[4.2.1 Strengths 34](#_Toc456104571)

[4.2.2 Limitations 34](#_Toc456104572)

[5 References 36](#_Toc456104573)

[6 Annex 1: synthetic estimation using Stata 43](#_Toc456104574)

[6.1 Synthetic estimates 43](#_Toc456104575)

[6.2 Calculating confidence intervals for prevalence estimates using boot strap procedures 47](#_Toc456104576)

[7 Annex 2: Stepwise forward and backward selection model 50](#_Toc456104577)

[8 Annex 3: ROC curves 52](#_Toc456104578)

[9 Annex 4: Detailed results for prevalence comparison 53](#_Toc456104579)

**HSfE 2014 depression prevalence model:**

**Technical Document**

# Background

Depression is a broad and heterogeneous disease. Central to its diagnosis is depressed mood and/or loss of pleasure in most activities. Severity of the disorder is determined by both the number and severity of symptoms, as well as the degree of functional impairment.[1 ,2] Depression worsens the health effects produced by other comorbidities. The effects of depression combined with other common chronic conditions (angina, arthritis, diabetes and asthma) has been observed to be worse than any other combination of these.[3]

Mental ill health presents a significant and complex public health problem; in the UK, mental ill health is the leading cause of disability, accounts for 28% of the national burden of disease and carries estimated economic costs of between £70-100 billion per year.[4] Direct costs of mental health in England are now around £22.5 billion a year – that includes spending in health and social care and a variety of other agencies, but not the indirect costs of the impact on the criminal justice system and in lost employment.[5] Within the NHS it accounts for more than 12 per cent of the total budget.[6] Current estimates suggest that the economic cost relating to mental illness is increasing, and that it is set to rise to more than £60 billion in England by 2026.{McCrone, 2008 #94} Since 2009 the number of working days lost to stress, anxiety and depression has increased by 24% and those lost to serious mental illness has doubled.{Office for National Statistics, 2014 #226} The Chief Medical Officer indicated in her 2013 annual report that over a similar time period, the real terms investment in mental health expenditure has reduced since 2011.{Thornicroft, 2013 #227} She stated that investment in mental health services and research is needed to strengthen the evidence base and services offered to improve the nation’s mental health.

The Adult Psychiatric Morbidity Survey (APMS) series highlights the inequalities that are present for mental illness across gender, age, ethnicity and varying income groups.[7] Besides observing inequalities in the prevalence of the symptoms but also in the diagnosis and treatment of mental illnesses. Mental illness has a significant negative impact on people’s lives. Complex relationships between mental health and physical health, specifically long term conditions, have been widely reported.[8] Long term conditions such as diabetes and cardiovascular diseases have been found to increase of one’s likelihood of experience mental health problems by two to three times more.[8] Among women in childbearing age, a study found a prevalence of 5.1% for antenatal depression and 13.3% for postnatal depression from 1994 to 2009 in the UK.[9]

Data from the new 2014 APMS was not available in time for this analysis. The most recent UK population prevalence data on depression comes from the Health Survey for England (HSfE) 2014, the data source we used for this analysis.[10] Chapter 2 of the published HSfE report is on mental health problems, and states that 26% of all adults reported having ever been diagnosed with at least one mental illness.{Bridges, 2015 #225} A further 18% of adults reported having experienced a mental illness but not having been diagnosed. Women were more likely than men to report ever having been diagnosed with a mental illness (33% compared with 19%). The most frequently reported mental illness ever diagnosed was depression, including post-natal depression, with 19% of adults (13% of men, 24% of women) reporting this. The next most frequently reported conditions ever diagnosed were panic attacks, mentioned by 8% of adults, and generalised anxiety disorder, mentioned by 6%.

Lifetime prevalence of each other condition was very low, at 3% or less.

## Diagnosis of depression

A formal diagnosis using the International Classification of Diseases version 10 (ICD-10) classification system requires at least four out of ten depressive symptoms,[11] whereas the Diagnostic and Statistical Manual of Mental Disorders version 5 (DSM-V) system requires at least five out of nine for a diagnosis of major depression.[12 ,13] Symptoms should be present for at least 2 weeks and each symptom should be present at sufficient severity for most of every day. Both diagnostic systems require at least one (DSM-V) or two (ICD‑10) key symptoms (low mood, loss of interest and pleasure or loss of energy) to be present.[14]

Increasingly, it is recognised that depressive symptoms below the DSM‑V and ICD‑10 threshold criteria can be distressing and disabling if persistent. Therefore updates cover 'subthreshold depressive symptoms', which fall below the criteria for major depression, and are defined as at least one key symptom of depression but with insufficient other symptoms and/or functional impairment to meet the criteria for full diagnosis. Symptoms are considered persistent if they continue despite active monitoring and/or low-intensity intervention, or have been present for a considerable time, typically several months. (For a diagnosis of dysthymia, symptoms should be present for at least 2 years.[14]

Diagnosis of hypertension frequently occurs in primary care. A meta-analysis by Mitchell et al of 41 studies that assessed the accuracy of unassisted diagnoses of depression by GPs found that they correctly identified depression in 47·3% (95% confidence intervals/CIs 41·7% to 53·0%) of cases and recorded depression in their notes in 33·6% (22·4% to 45·7%), although recording will currently be much higher in UK primary care electronic health records (EHRs). [15] Of these 19 studies assessed both rule-in and rule-out accuracy; from these studies, the weighted sensitivity was 50·1% (41·3% to 59·0%) and specificity was 81·3% (74·5% to 87·3%). At a rate of 21·9%, the positive predictive value was 42·0% (39·6% to 44·3%) and the negative predictive value was 85·8% (84·8% to 86·7%). This finding suggests that for every 100 unselected cases seen in primary care, there are more false positives (n=15) than either missed (n=10) or identified cases (n=10). Accuracy was improved with prospective examination over an extended period (3–12 months) rather than relying on a one-off assessment or case-note records.

GPs can rule out depression in most people who are not depressed; however, the modest prevalence of depression in primary care found by Mitchell et al means that misidentifications outnumbered missed cases. Diagnosis could be improved by re-assessment of individuals who might have depression. In this setting, it is estimated that about 84.4% to 86.7% of depression cases are identified, with the number of patients erroneously identified as depressed (95% CI 55.7-60.4%) actually outnumbering the missed cases.

The current situation in UK primary care is likely to be considerably better than the meta-analysis. Depression has been part of the Quality & Outcomes Framework (QOF) GP pay for performance scheme for several years.[16] MH001, the first QOF depression indicator, requires that practices “establish and maintain a register of patients with schizophrenia, bipolar affective disorder and other psychoses and other patients on lithium therapy”. Unlike some chronic diseases, remissions of depression spontaneously or in response to treatment are common. QOF remission codes can be used if the patient has been in remission for at least five years, that is where there is:

* no record of anti-psychotic medication
* no mental health in-patient episodes; and
* no secondary or community care mental health follow-up for at least five years.

Where a patient is recorded as being ‘in remission’ they remain on the MH001 register (in case their condition relapses at a later date), but they are excluded from the denominator for other related mental health indicators.

## Types of depression

Prevalence estimates from various studies are shown in Table 1. Within the overall definition there are several important categories of depression. These are discussed below.

### Major depressive disorder

The 12-month prevalence of “major” depressive disorder is 6.6% and the lifetime prevalence is 16.2%. [17] Most age-of-onset distributions propose that major depressive disorder is prevalent for the entire lifetime. [18] Major depression can cause deterioration in health which is equivalent to that caused by other chronic diseases, for example, angina, arthritis, asthma, and diabetes, and furthermore, when depression occurs together with these comorbidities, it results in significantly greater decrements in health, than when the diseases occur alone.[19] Therefore, it is important that primary care providers do not ignore the presence of depression in patients who also have a chronic physical illness.[17]

When ensuring the appropriate diagnosis and management of clinical depression, it is important to take be aware of over-detection and under-detection.[20] Whilst a meta-analysis found that general practitioners exclude depression correctly in most people who do not have clinical depression, false positives can outnumber missed cases.[21]

There is a wide variation in the prevalence of depression between countries and also across studies which are conducted in the same country.[22] This has been argued by some that these disparities in prevalence may arise due to methodological issues, for example, recall bias and the age of the subjects being interviewed.[22]

### Bipolar disorder

According to population surveys (diagnostic interviews with standardised criteria), the lifetime prevalence of bipolar disorders is estimated to be around 0.8% for bipolar I and 1.1% for bipolar II.[23],[24] Although the prevalence is low, bipolar disorder is a leading cause of premature mortality as a result of suicide and related medical illnesses, for example, diabetes and cardiovascular disease.[25] The true prevalence of bipolar disorder may be underestimated as there is inconsistency between the rates of bipolar disorder estimated in large-scale community surveys [23] and from prospective longitudinal studies.[26] Prospective studies suggest that the inclusion criteria of symptoms and the diagnostic threshold for bipolar disorder are too restrictive to detect bipolar disorder in the general population, thus the burden of bipolar disorder may be greater than the present estimates.[27]

Table 1 Prevalence estimates of depression from literature

| Publication[[1]](#footnote-1) | Country | Type of study | Years covered | Total\* | Male\* (%) | Female\* (%) | Age group (years) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mavreas *et al, 1986* | Greece | Cross-sectional design | 1984 | \_ | 4.3 | 10.2 | All ages |
| Sandanger *et al,* 1999 | Norway |  | 1989-1991 | \_ | 0.7  0.5  1.2 | 3.1  4.4  7.9 | 20-39  40-59  60-79 |
| Jenkins *et al, 1997* | UK | Cross-sectional design | 1997 | \_ | 1.7  2.0 | 2.7  1.1 | 16-54  54-64 |
| Bener *et al,* 2012[28] | Qatar | Prospective cross-sectional design | 2009 | N= 224 | 36.2  41.0  22.9 | 48.7  43.7  7.6 | 18-34  35-49  50+ |
| Busch *et al,* 2013[29] | Germany | Cross-sectional design | 2013 | 9.9 (7.8-12.3)  7.9 (5.7-10.8)  8.4 (6.7-10.5)  8.2 (6.7-10.1)  7.2 (5.6-9.2)  6.1 (4.3-8.6) | 8.0 (5.5-11.5)  5.3 (5.5-11.5)  7.0 (5.0-9.7)  6.1 (4.5-8.3)  4.5 (2.9-6.9)  4.2 (2.6-7.0) | 11.8 (9.0-15.3)  10.5 (7.3-14.7)  9.9 (7.6-12.8)  10.4 (8.1-13.3)  9.8 (7.1-13.4)  7.7 (5.1-11.4) | 18-29  30-39  40-49  50-59  60-69  70-79 |
| Li *et al,* 2014[30] | China | Meta-analysis | 1987-2012 | 23.6 (20.3-27.2-) | 15.7 (12.1-21.0) | 22.2 (17.7-27.4) | All ages |
| Bijl *et al,* 1998 | Netherlands | Cross-sectional | 1998 | \_ | 1.9 | 3.4 | 18-64 |
| Vazquez-Barquero *et al,* 1987[31] | Spain | Cross-sectional | 1984 | \_ | 4.3 | 5.5 | >17 |
| *Park et al,* 2012[32] | South Korea | Cross-sectional | 2008 | 27.8 (26.7-29.0) | 22.4 (20.8-24.1) | 31.1 (29.6-32.7) | All ages |
| Fuhrer *et al,* 2003 | France | Prospective population-based cohort | 1988-1996 | \_ | 6.4 | 8.4 | 75+ |
| Murata *et al,* 2008 | Japan | Cross-sectional | 2003 | 35.5 | 34.4 | 36.5 | 75-79 |

Table 2 shows the prevalence of depression by age and sex in England in 2000, from the study by Thomas et al. However the most recent published data is from the HSfE 2104, which is shown in Table 3.{Bridges, 2015 #225} The most frequently reported mental illness ever diagnosed was depression, including post-natal depression, with 19% of adults (13% of men, 24% of women) reporting this.

Table : Prevalence of depression by age and sex in England in 2000[6]

| Age (years) | Cases per 1000 patients | | Number of cases | | Total |
| --- | --- | --- | --- | --- | --- |
|  | **Male** | **Female** | **Male** | **Female** |  |
| 16-24 | 15.6 | 45.2 | 48,156 | 131,939 | 180,094 |
| 25-34 | 31.5 | 87.1 | 121,555 | 317,654 | 439,209 |
| 35-44 | 40.3 | 102.7 | 154,829 | 381,387 | 536,215 |
| 45-54 | 48.0 | 108.9 | 157,555 | 357,649 | 515,205 |
| 55-64 | 48.5 | 102.8 | 122,458 | 265,964 | 388,422 |
| 65-74 | 40.1 | 98.6 | 76,635 | 214,632 | 291,268 |
| 75-84 | 49.8 | 104.9 | 53,510 | 172,477 | 225,987 |
| 85+ | 56.5 | 96.5 | 15,261 | 69,808 | 85,069 |
| Total | 29 | 70.1 | 749,958 | 1,911,510 | 2,661,468 |

Table : percent prevalence of diagnosed mental illness, by age group, HSfE 2014

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 16-24 | 25-34 | 35-44 | | 45-54 | | 55-64 | | 65-74 | | 75-84 | | 85+ | |  | |
| Depression (including post-natal) | 11 | 17 | 21 | 22 | | 25 | | 19 | | 12 | | 7 | | 19 | |
| Panic attacks | 4 | 11 | 8 | 10 | | 10 | | 6 | | 5 | | 2 | | 8 | |
| Generalised anxiety disorder | 4 | 7 | 6 | 5 | | 8 | | 4 | | 3 | | 4 | | 6 | |
| Post-traumatic stress | 1 | 2 | 2 | 2 | | 3 | | 1 | | 1 | | - | | 2 | |
| Phobia | 1 | 1 | 1 | 1 | | 1 | | 1 | | 0 | | 1 | | 1 | |
| Obsessive compulsive disorder | 2 | 2 | 1 | 1 | | 0 | | 0 | | - | | - | | 1 | |
| Nervous breakdown | 1 | 1 | 2 | 3 | | 4 | | 3 | | 4 | | 3 | | 3 | |
| Eating disorder | 2 | 1 | 2 | 1 | | 1 | | 0 | | 1 | | - | | 1 | |
| Bipolar disorder | 0 | 1 | 1 | 1 | | 0 | | 0 | | - | | 1 | | 1 | |
| Personality disorder | 1 | 0 | 0 | 1 | | 1 | | - | | - | | 1 | | 1 | |
| Psychosis or schizophrenia | 0 | 1 | 0 | 1 | | 0 | | 0 | | - | | - | | 1 | |
| Other mental, emotional or neurological problem/condition | 1 | 1 | 0 | 1 | | 2 | | 0 | | 1 | | 2 | | 1 | |
| Seasonal affective disorder | 0 | 1 | 0 | 1 | | 2 | | 1 | | 0 | | - | | 1 | |
| ADHD or ADDc | 2 | 1 | 0 | 0 | | 0 | | 0 | | - | | - | | 1 | |
| Dementia | - | - | - | - | | 0 | | 0 | | 1 | | 2 | | 0 | |
| Alcohol or drug dependence | 1 | 2 | 2 | 1 | | 2 | | 0 | | - | | - | | 1 | |

## Risk factors

For this project, we conducted a non-systematic literature search on the different risk factors affecting depression according to the most recent literature available (Table 4). We have summarized the evidence from different sources in this section, by risk factor, as well as the main issues identified.

Table : Hypertension risk factor list

| Risk factor | References |
| --- | --- |
| Sex | [33], [34], [35] |
| Age | [35] |
| Ethnic group | [36], [37] |
| Obesity/BMI | [38], [39] |
| Socioeconomic status/Financial problems | [35] |
| Education | [35], [40] |
| Civil status | [41] |
| Anxiety | [42], [43] |
| Long-lasting illness | [35] |
| Physical activity | [44], [45], [46] |
| Alcohol intake | [47], [43] |
| Drug abuse | [43] |
| Sleep disorders | [48] |

### Sex

Historically women are more likely to report depressive symptoms than men. Although the odds ratios of both genders vary between some countries, part of this variation might be explained by the method used to define these symptoms. A study from 2010 looking into three different cohorts with participants from the UK, EU and the USA separately found that women from each study were 1.78, 2.35 and 1.5 times more likely than male participants to report depressive symptoms. [33]

### Age

Increasing age has been associated with an increased risk of depression. A study examining data on symptoms of major depressive episodes from the Alameda County Study found that individuals who were 70 to 79 years of age were at 1.63 times greater odds (95% CIs 1.13-2.37) of having depression compared to adults aged 50 to 59 years.[35] Adults of 80 years or more were twice more likely (95% CIs 1.23-3.24) to have depression compared to 50 to 59 year olds. However, there was no significant difference observed in the odds of having depression in 60 to 69 year olds and 50 to 59 year olds. In this study, depression was measured using the DSM-IV diagnostic criteria for major depressive episodes.

### Ethnic group

Evidence from qualitative research suggests that perceptions of depression, including its consequences and chronic nature, might differ between women from Black African and White ethnic backgrounds in the UK, potentially affecting the lower detection rates observed in the former. [36] A study using nationally representative data from the Collaborative Psychiatric Epidemiology Studies in America found African-Americans were at a significantly lower risk of meeting the criteria for any past-year depressive disorder compared to non-Latino white individuals.[37]

### Body mass index (BMI)

Evidence of the effects of BMI on depression is mixed. In a systematic review evaluating this relationship, Atlantis et al found four cohort studies supporting the hypothesis that obesity is a risk factor for depression. However, evidence from the cross-sectional studies reviewed was limited to women in the US, and was not found in men in other countries.[38] According to Luppino et al, some of the risk factors for metabolic syndromes such as obesity could also have an effect on depression, which would also increase the risk for developing obesity.[39] A cohort study with a 5 year follow-up found individuals with a BMI of 30kg/m2 or more were 1.79 times more likely to develop major depression compared to those with a BMI less than 30kg/m2.[49]

### Education

According to the results observed in a cohort (N=33.774) from the HUNT study, respondents with lower educational levels were at higher risk of suffering depression compared to those with higher levels. An interaction with age was also observed, although the direction of such is not clear possibly due to a cohort effect.[40]

### Marital status

According to a study using the Canadian National Population Health survey (1994-2004), there is a bi-directional correlation between major depression and separated or divorced status, each increasing the chances that the other takes place. [41]

### Socioeconomic status

The Alameda County Study found that individuals who were having financial problems were at a 2.53 times greater odds of having depression in comparison to those without any financial problems.[35]

### Anxiety

Major depressive disorder was assumed to precede generalised anxiety disorder until a 32-year prospective follow-up study challenged this notion. Indeed, the reverse pattern seems to be frequently present, and the combination of generalised anxiety disorder and major depression might represent an additional burden. Social anxiety disorder (social phobia) is now also regarded as an important and consistent risk factor for the development of severe depression.[42]

### Long-lasting illness

In addition, differences between white and south Asian ethnic groups in the UK have been observed in the odds of being diagnosed with depression, however, this might be more related to differences in the diagnosis of the condition rather than the probability of suffering it.[50] Similarly, a longitudinal study conducted in the USA among 1034 individuals observed that white respondents with more chronic conditions were more likely to develop depressive symptoms and vice versa, however, this relation was not found in Black respondents.

### Physical activity

Although some studies have observed some protective effect between low and moderate physical activity and depression, evidence is still limited in this area. A follow-up study including 6661 individuals from the HUNT study found some evidence of a protective effect among men performing low physical activity and women performing medium physical activity compared to inactive individuals.[45] A more recent cross-sectional study of individuals recruited in the HUNT study found a correlation between reduced chances of suffering depression and frequency, duration and intensity of physical exercise among men and women.[44] Moreover, a 20-year cohort study since childhood found that males who were persistently active had a 0.35 times reduced risk of having depression compared to those who were persistently inactive, showing the protective effect of physical exercise. [46] However, this relationship was not observed for the female participants.

### Alcohol abuse

In a recent cohort study of 1034 respondents in the US, it was observed that the risk of developing depressive symptoms was higher for individuals with more chronic medical conditions at baseline. However this correlation was only observed for White respondents but not African-Americans.[47] Moreover, other studies have found that the risk for developing depression varies between individuals with different conditions (such as Type 2 diabetes) also depending on their ethnic background, being higher in individuals of White ethnicity compared to South Asians.[50]

### Sleep disorders

Insomnia and depression have a bidirectional relationship.[51] Cross-sectional studies have shown that there is a significant relationship between depressive symptoms and insomnia, with insomnia longitudinally linked with the development of depression and poor outcomes of treatment.[51] A cohort study using data from two general health surveys of the adult population in Norway found a significant association between insomnia and the risk of depression.[48] Adults with insomnia were 1.8 times more likely to have depression compared to those without insomnia.

Table 5 summarises these findings.

Table : Risk factors and their ORs from various studies

| Risk factor | Type of Odds Ratio | Odds Ratio | 95% CI | Effect on Outcome |
| --- | --- | --- | --- | --- |
| Sex |  |  |  |  |
| Male | Adjusted Hazard Ratio [33] | 1.00 |  | Reference |
| Female | Adjusted Hazard Ratio [33]  Odds Ratio [35] | 1.78  1.63 | 1.61-1.98  1.20-2.72 | Risk factor  Risk factor |
| Age (years) |  |  |  |  |
| 50-59 | Odds Ratio [35] | 1.00 |  | Reference |
| 60-69 | Odds Ratio [35] | 1.02 | 0.69-1.51 | Not significant |
| 70-79 | Odds Ratio [35] | 1.63 | 1.13-2.37 | Risk factor |
| 80+ | Odds Ratio [35] | 2.00 | 1.23-3.24 | Risk factor |
| Ethnicity |  |  |  |  |
| Non-Latino White | Weighted Odds Ratio [37] | 1.00 |  | Reference |
| Latino | Weighted Odds Ratio [37] | 1.19 | 0.85-1.66 | Not significant |
| Asian | Weighted Odds Ratio [37] | 0.67 | 0.41-1.10 | Not significant |
| Afro-Caribbean | Weighted Odds Ratio [37] | 1.33 | 0.56-3.14 | Not significant |
| African American | Weighted Odds Ratio [37] | 0.63 | 0.48-0.82 | Reduced risk |
| Body Mass index (BMI)/kg/m2 |  |  |  |  |
| <30 | Odds Ratio adjusted for age, gender, education, marital status, social support, physical health problems [49] | 1.00 |  | Reference |
| ≥30 | Odds Ratio adjusted for age, gender, education, marital status, social support, physical health problems [49] | 1.79 | 1.06-3.02 | Risk factor |
| Education |  |  |  |  |
| 12+ years | Odds Ratio [35] | 1.00 |  | Reference |
| <12 years | Odds Ratio [35] | 2.05 | 1.45-2.89 | Risk factor |
| Civil status |  |  |  |  |
| Married | Adjusted Hazard ratio [41] |  |  | Reference |
| Separation/divorce | Adjusted Hazard ratio [41] | 1.3 | 1.0-1.5 | Risk Factor |
| Married | Odds Ratio [35] | 1.00 |  | Reference |
| Divorced, separated, or widowed | Odds Ratio [35] | 1.24 | 0.90-1.70 | Not significant |
| Financial problems |  |  |  |  |
| None | Odds Ratio [35] | 1.00 |  | Reference |
| Any | Odds Ratio [35] | 2.53 | 1.83-3.51 | Risk factor |
| Neighborhood quality |  |  |  |  |
| No problems | Odds Ratio [35] | 1.00 |  | Reference |
| Some problems | Odds Ratio [35] | 1.61 | 1.10-2.33 | Risk factor |
| Serious problems | Odds Ratio [35] | 3.24 | 2.28-4.59 | Risk factor |
| Social isolation |  |  |  |  |
| Low | Odds Ratio [35] | 1.00 |  | Reference |
| Medium | Odds Ratio [35] | 1.55 | 1.05-2.28 | Risk factor |
| High | Odds Ratio [35] | 3.07 | 2.15-4.40 | Risk factor |
| Physical activity |  |  |  |  |
| Females |  |  |  |  |
| Persistently inactive | Relative Risk adjusted for self-reported health status, language spoken at home in childhood, no. of live births (females only), current occupation, age, depressed mood [46] | 1.00 |  | Reference |
| Decreasing | Adjusted Relative Risk [46] | 1.04 | 0.58-1.85 | Not Significant |
| Increasing | Adjusted Relative Risk [46] | 0.73 | 0.39-1.38 | Not significant |
| Persistently active | Adjusted Relative Risk [46] | 0.62 | 0.35-1.11 | Not significant |
| Males |  |  |  |  |
| Persistently inactive | Adjusted Relative Risk [46] | 1.00 |  | Reference |
| Decreasing | Adjusted Relative Risk [46] | 0.40 | 0.15-1.05 | Not significant |
| Increasing | Adjusted Relative Risk [46] | 0.31 | 0.11-0.92 | Reduces risk |
| Persistently active | Adjusted Relative Risk [46] | 0.35 | 0.15-0.81 | Reduces risk |
| Alcohol and Drugs |  |  |  |  |
| Alcohol abuse | Adjusted OR [43] | 3.6 | 2.6-5.0 | Risk factor |
| Alcohol abuse with dependence | Adjusted OR [43] | 6.1 | 3.7-10.0 | Risk factor |
| Drug abuse | Adjusted OR [43] | 4.6 | 2.9-7.5 | Risk factor |
| Drug abuse with dependence | Adjusted OR [43] | 10.4 | 4.7-23.0 | Risk factor |
| Anxiety disorder | Adjusted OR [43] | 9.4 | 8.1-11.0 | Risk factor |
| Chronic medical condition |  |  |  |  |
| None | Odds Ratio [35] | 1.00 |  | Reference |
| 1 | Odds Ratio [35] | 2.05 | 1.35-3.11 | Risk factor |
| 2+ | Odds Ratio [35] | 4.47 | 3.06-6.52 | Risk factor |
| Sleep disorders | Adjusted Odd Ratio [48] | 1.8 | 1.6-2.0 | Risk factor |

# Methods

## Data source & sampling

The depression model described here uses data from the 2014 Health Survey for England (HSfE).[10] New topics for that year include hearing and mental health. The achieved sample for the 2014 survey was 8,077 adults (aged 16 and over) and 2,003 children (aged 0-15). The 2014 data consists of a household and an individual level files. The latter, hse2014ai (10,080 records) contains data for all individuals in the General Population Sample in co-operating households who gave a full interview, and was used for this project. HSfE used a stratified probability sampling design, with two stage sampling. In the first stage, a random sample of primary sampling units (PSUs) based on postcode sectors was selected. Within each PSU, a random sample of 16 postal addresses of private households was selected. All adults aged 16 years and over at each household were selected for the interview (up to a maximum of ten adults per household). Nurse visits were offered to all participants who were interviewed. In these visits, nurses questions asked about self-reported and doctor-diagnosed mental health conditions, including depression, post-natal depression and bipolar disorder.[10] A final population of individuals over 16 years old, with a nurse visit and who answered the mental health questionnaire (n=5,485) was selected for this study.

### Outcome variables

For this model we investigated two different definitions for depression, diagnosed and undiagnosed. The first outcome, patient-reported doctor-diagnosed depression, was derived from the number of patients that reported being diagnosed with depression by a health professional (variable BCond\_7\_HProf). The actual question is: *“Did a doctor, psychiatrist or other professional tell you that you had the following condition: Depression?”* No time restriction was put on the definition, so the diagnosis may have been made years ago and the disease will in many cases have resolved. The second outcome, undiagnosed depression, was derived from those respondents that reported they considered they experienced depression (variable DiagCX7), but were not diagnosed by a health professional. The actual question is: *“Do you think that you have ever experienced any of these? Depression*”. There is also a question: *“In the last 12 months have you had: Depression*?” which could be used to investigate annual incidence/point prevalence.

We decided that, although the first question could be used to measure the extent of undiagnosed depression, use of the second question i.e. health professional diagnosed depression was more consistent with the outcome definitions for other models, and best for gauging the effectiveness of case-finding by practices. Another option might be to sum or in other ways combine the two definitions to get the best estimate of overall population prevalence at any point in time.

### Risk factors

We used the literature review described in the Background to extract HSfE data on risk factors. There were two main reasons why some risk factors from the literature were not used in the final model. Firstly, the data was not available in HSfE (for example information on sleeping disorders) or some form of data was available, but did not adequately fit the needs of the model. More importantly, a number of variables were excluded from the model due to the limited number of respondents. This is the case for alcohol and drug abuse or self-reported diagnosed anxiety. Unlike the HSfE hypertension model, it was not possible to merge data from several years for the depression model, as 2014 was the first year in which depression and other mental health conditions are included with the required format.

### Missing values

There were a number of missing values in HSfE, identified as respondents who either did not know or refused to provide an answer to a specific question. In HSfE, all these values are coded with a different negative number depending on the reason why the answer was not obtained. In order to deal with missing values, multiple imputation was used. Given the evident problems that complete-case analysis brings due to the deletion of incomplete observations, multiple imputation is a widely accepted technique to estimate those values missing.[52-55] Under the assumption that missing values on the dataset are missing at random (MAR), regression analysis is conducted on variables registered as imputed (i.e. with missing values) using other complete variables. Through this analysis, missing values were inferred from the relation of the imputed variable with the other complete variables. This process will be conducted a minimum of five times, a mean of these n=5 analyses will be the imputed value substituting the missing information on each observation.

## Regression modelling

All statistical analysis was carried out in Stata SE-14. The choice of variables for original inclusion in the merged dataset included all those known to be depression risk factors. The HSfE dataset has a nested or hierarchical structure so three variables related to the sampling strata were included: area (sample point), cluster (stratification level), and wt\_nurse (weights accounting for non-respodents and survey structure). These were used in the regression models to adjust for clustering of respondents.

We fitted univariate then multivariate logistic regression models for diagnosed depression to produce odds ratios (ORs) and regression coefficients. Model specification was conducted using forward stepwise selection and each newly introduced variable was tested using Wald tests. A range of multivariate regression models were fitted in order to obtain the best performing.

The modelling and estimation of the effects of interest was carried out using the logit command. The initial output consisted of two tables: one with the estimated regression coefficients, corresponding p-values and 95% confidence intervals, and another with the estimated odds ratios (exp(b)), which in the table appear as relative risk ratios (RRRs) and 95% confidence intervals. A positive sign of the estimated coefficient is associated with an increase in the odds of the outcome i.e. had diagnosed depression, and a negative sign is associated with a decrease in the odds. Since Prob (A) = Odds (A) / 1+ Odds (A), for uncommon outcomes such as high blood pressure, RRR can be assumed to be the same as the odds ratio (OR).

For categorical variables the effects are estimated relative to the reference category. Stata uses the first category as reference (baseline OR). Separate baseline odds were estimated for each gender, and also according to ethnicity, age band, area-based deprivation score etc. The model can be used to derive the prevalence ratios for depression for subjects with various combinations of risk factors in relation to baseline. The prevalence in each age group, gender, ethnic group, area of residence and level of deprivation category were derived from the odds, using the formula: prevalence = odds/(1 + odds).

There is an interaction between the effects of two exposures if the effect of one exposure varies according to the level of the other exposure.[56] For example, there might be an interaction between the depression risk factors of education level and social class. An alternative term for interaction its effect modification. In this example, we can think of this as educational level modifying the effect of social class. For this model, we tested for interactions between HSfE predictor variables for risk factors. Wald tests were used to compare the possible interactions and decide whether they should be included.

## Internal validation

Ideally the best prediction should result from utilising the most information in the regression model. However only a limited range of HSfE variable data is either available or can be estimated at the small population level. Therefore we validated a local model (that only used locally available data) by comparing it, in terms of prediction, to a complete model including all available and significant HSfE variables. In addition we internally validated the models by generating receiver operating characteristic (ROC) curves, by using the ***predict*** regression post-estimation command to generate for each HSfE informant the probability of having depression using the derived odds ratios (ORs), and by using these probabilities to examine sensitivity and specificity.

## Local prevalence estimates

### Data sources

Various sources were used for local risk factor data as follows:

* practice age/sex breakdowns were obtained from the Health & Social Care Information website: <http://www.hscic.gov.uk/searchcatalogue?topics=1%2fPrimary+care+services%2fGeneral+practice&sort=Most+recent&size=10&page=2#top>
* Top qualification from Census 2011 data available on the Neighbourhood Statistics website, mapped to practices: <http://www.neighbourhood.statistics.gov.uk/dissemination/Download1.do?&nsjs=true&nsck=false&nssvg=false&nswid=1920>
* Ethnicity: from Census 2011 data available on the Neighbourhood Statistics website, mapped to practices: <http://www.neighbourhood.statistics.gov.uk/dissemination/datasetList.do?$ph=60&updateRequired=true&step=1&CurrentTreeIndex=-1&Expand10=1#10>
* Employment category from Census 2011 data available on the Neighbourhood Statistics website, mapped to practices: <http://www.neighbourhood.statistics.gov.uk/dissemination/Download1.do?&nsjs=true&nsck=false&nssvg=false&nswid=1920>
* Index of Multiple Deprivation (IMD) mapped to practices: Department of Communities & Local Government: <https://www.gov.uk/government/collections/english-indices-of-deprivation>. IMD quintiles are not available nationally for any level, so we had to calculate the percentage of people in each quintile:
  + we divided Lower Super Output Areas (LSOAs) into quintiles according to their IMD score (using 2013-2014 criteria)
  + we converted them into populations, using the number of people contributed by each LSOA to each practice
  + we summed up for each practice IMD scores all the people in different LSOAs for each quintile
  + we turned these into percentage of the practice population in each IMDS quintiles.
* Body Mass Index: Active People Survey (APS) commissioned by Sport England: obtained from PHE National Obesity Observatory website: <http://www.noo.org.uk/LA/obesity_prev/adults>
* Limiting long-lasting illness from Census 2011 data available on the Neighbourhood Statistics website, mapped to practices: <http://www.neighbourhood.statistics.gov.uk/dissemination/Download1.do?&nsjs=true&nsck=false&nssvg=false&nswid=1920>

### Small population estimation methods

Derived ORs (or rather, regression coefficients) are used to estimate prevalence in small population subgroups. Local population breakdowns for each risk factor are used, where these are available. ICL has a wide range of small population risk factor prevalence breakdowns, including age, sex, deprivation, alcohol consumption, ethnicity, long-term illness, anxiety, BMI and marital status. The local model uses locally available data and so includes only those variables that are available at local population level i.e. age, sex, socioeconomic status, BMI, and other disease conditions. The steps in applying the prevalence estimates are as follows and in the equations below:

* Use the regression coefficients to generate log odds (since they are from a logistic regression model) for each risk factor subcategory
* Generate a similar table of odds by exponentiation
* Generate a similar table of prevalence in each risk factor subcategory using the epidemiologic formula
* Produce a matching table of small population subcategories. If there are no corresponding local data with a sufficiently granular breakdown e.g. ethnicity by age by sex, this requires deciding how each risk factor should be attributed across other risk factor categories, with evenly as the default. For example, we used the national age/sex/ethnicity breakdown from the Census and age breakdowns from the HSfE to attribute this data at small population levels. The actual breakdown will be somewhat different and needs to be borne in mind as another source of potential error.
* Multiply the population cells by the corresponding prevalence to estimate the number of people in each cell with the disease

In mathematical notation:

Predicted log odds of prevalence = *b0* + *b1x1i* +  *b2x2 i* + *b3x3 i* +  *b4x4 I*

where *b0* = regression constant, *b1, b2,  b3, b4*= other regression coefficients

*x 1 i, x2 i, x3 i, x4 i* = value of risk factors for individual ***i***

(NB since all the variables are binary variables, *x* =1 if specified risk factor is present, *x*=0 if it is absent). Predicted log odds of prevalence for a community of *n* individuals is derived by averaging over the values for all individuals included in the community:

Predicted log odds of prevalence in community of *n* individuals:

= 1/*n* ∑i=1n (*b0* + *b1x1i* +  *b2x2 i* + *b3x3 i* +  *b4x4 i)*

= *b0* + *b1p1* +  *b2p2* + *b3p3* +  *b4pp4*

where *p1 , p2, p3, p4*=proportion of individuals in the community with characteristic *x1 , x2 , x3 , x4 .* (i.e. proportion with *x.*=1 rather than *x.*=0 as in the remainder).

The predicted prevalence for an individual is derived from their predictive log odds using:

prevalence = exp(log odds)/[1+exp(log odds)]

=*exp*(*b0* + *b1x1i* +  *b2x2 i* + *b3x3 i* +  *b4x4 i) /[1+* *exp*(*b0* + *b1x1i* +  *b2x2 i* + *b3x3 i* +  *b4x4 i)]*

Predicted prevalence in community of *n* individuals:

= 1/*n* ∑i=1n{*exp*(*b0* +*b1x1i* +*b2x2 i* +*b3x3 i* +*b4x4 i)/[1+* *exp*(*b0* +*b1x1i* +*b2x2 i* +*b3x3 i* +*b4x4 i)]*

Unfortunately, the equation above does not simplify to a linear combination of the predictor variables (in the way the mean log odds does). The average/overall prevalence is not the same as the prevalence for a person with “average” risk factors. So, for instance, it cannot be found by taking exp(log odds)/[1+ exp(log odds)] of the average log odds. There is no linear relationship with the regression coefficients, and with proportions of population with specified risk factors.

In order to find a synthetic estimate of prevalence, ideally we need to know the joint distributions of the included risk factors in the relevant population (the population on which are synthetic estimates are required). Ideally, we would know how many people in the population have each specific combination of risk factors. In practice, it might be good enough to know the distribution of some risk factors individually, rather than in combination. For instance, we might know what proportion of the population are smokers, and what proportion are ex-smokers, but not how many smokers we have by age and sex. In this situation, we have assumed that the same proportion of all ages and both genders are smokers and ex-smokers. Even if this is not exactly correct, then the synthetic estimate of prevalence may still be a reasonably accurate estimate (assuming that the smoking distribution does not vary too much by age, sex and other included risk factors). This is considered a good enough approach, and the best possible based on the information currently available in many cases.

In practice, we know the population distributions by age and sex, therefore we do not need to make the assumption that the proportion of males is the same for each age group. We use the more precise method of using the actual proportions of males in each age group. From the ELSA longitudinal survey we also know that older people/ older females in particular are generally less educated (on the basis of qualifications held). Therefore we apply the proportions with any educational qualifications according to age and sex group.

For other risk factors, we do not know whether these risk factors are more or less common in males than in females, nor according to age group, nor educational status i.e. we do not know their distributions in combination with any of the other risk factors included in the model. Therefore we make the assumption that the distribution of all other risk factors (apart from afore-mentioned age, sex and educational status), is equal across all other risk factors. This makes the calculations somewhat easier, even though this assumption might make for slightly less accurate estimates, the loss of accuracy is not thought to be great.

In order to find the estimated prevalence for each population, it is necessary to calculate the synthetic prevalence of risk factors for each possible combination of risk factor (as included in the chosen disease-specific logistic regression model). The estimated prevalence for a population is then the weighted average of the prevalence estimates for each combination of risk factors, according to the estimated number of people with each risk factor combination in the population (the population on which synthetic estimates are sought). These calculations can be carried out in Excel (using VBA code to link prevalence and risk factor spreadsheets with formulae in a workbook) or in Stata software to produce confidence intervals as well as the estimates. We have developed two methods of producing these estimates, with CIs, in Stata. The first uses a “bootstrap” procedure, the second using inverse probability weighting. Ideally we wish to use both methods as an additional internal validation. However the short timeframe did not permit this so the estimates have been produced using inverse probability weighting only. More information about the bootstrap method is included in Annex 1.

### Local prevalence estimates: bootstrap method for local prevalence estimates

In summary, within Stata, a new set of variables is created, one for each combination of these risk factors pertinent to the logistic regression model for the chosen disease. With our dataset set up in this way, we can now use Stata’s “*predict*” command to give us the predicted log odds. Then we find the weighted average of these, averaged across all possible combinations of risk factors, using the weights calculated as above (stored in variable named xyz). The weighted average can be found using the “collapse” command, which results in one line of data per practice or MLSOA (using the population identifier as the *by* variable) in Stata.

We have also calculated in Stata CIs for prevalence estimates using a “bootstrap” procedure. There is uncertainty in these synthetic estimates of prevalence based on the imprecision not in the more usual sample of people from the population (since the estimates are not a sample but are externally applied), but in the estimated coefficients from the logistic regression equations. A bootstrap procedure can be used to construct confidence intervals on these synthetic estimates of prevalence, based on the imprecision in these logistic regression coefficients.

The philosophy underlying the bootstrap procedure is to consider that the people included in the data set used to derive the logistic regression equation represent the whole population of possible people. However, the whole population is effectively considered to contain thousands of copies of each of these people. Bootstrap samples are taken randomly from our initial populations (the subsets of the HSfE population that has complete data on appropriate risk factors). Logistic regression of the same risk factors can then be applied to this boot strap sample, i.e. we rerun the logistic regression that gave us our chosen predictive model. However, we get slightly different regression coefficients, because of the modified sample. Prevalence estimates are then derived for each combination of risk factors, based on these new regression equations.

This process is repeated 1,000 times, to find 1,000 different boot strap samples, by random sampling processes, and to then fit logistic regression equations on each. The prevalence estimates are calculated for each combination of risk factors, for each of these 1,000 boot strap samples. For each small population, a synthetic estimate is calculated for each boot strap sample, by appropriately weighting the prevalence estimates on each combination of risk factors (with the same weights as described above which reflect the anticipated prevalence of each combination of risk factors in the population). From these 1,000 synthetic estimates of prevalence of each population, a 95% confidence interval is calculated as the 2.5th to 97.5th centiles. Given that the estimates are distributed normally, these are taken to be mean +/- 1.96 SD (taking mean and SD of the 1,000 boot strap synthetic prevalence estimates for each specified region).

### Local prevalence estimates: inverse probability weights method

Inverse probability weighting methods are used to standardize from a sampled population to a target population. They are usually defined as a function of a panel of one or more sampling-probability predictor variables. For each combination of the predictor variables, the sampling probability weight is the ratio of the frequency of that combination in the target population to the frequency of that combination in the sampled population. Inverse probability weighting is therefore a generalization of direct standardization. In Stata, it is implemented by using a *pweight* qualifier on an estimation command. This normally implies the use of a Huber variance formula to generate the confidence limits.

The “local” model includes only those variables that are available at local population level. Local population breakdowns for each risk factor are used, where these are available. ICL has a wide range of small population risk factor prevalence breakdowns, including age, sex, deprivation, education levels, employment category, BMI, long-lasting illness, ethnicity. The local model uses locally available data.

In Stata, we may apply inverse-probability weights to a regression model, using the *pweight* qualifier to standardize the regression results from a sampled population to a target population. Alternatively, once we have regression results from one set of data, we may use a *pweight* qualifier with a post-estimation command (such as margins or the add-on packages *margprev* or *marglmean* [57 ,58] to compute a predicted mean or prevalence, using out-of-sample prediction. As an example, we may compute predicted prevalence for practices, using *margprev* in a scenario dataset for each practice.

To do this, we save the estimation results for the regression model in a .ster file (containing Stata estimation results). For each practice, we create a scenario dataset, with 1 observation for each possible combination of predictor values in that dataset. The sampling probability weights are the presumed probabilities of each combination of predictor values in that practice. In practice, we do not know the combination probabilities for each practice, as we only have estimates of marginal probabilities. (Such as a distribution of gender/age, a distribution of smoking status, and a distribution of ethnic origin, without any combination or joint distribution.) So, we assume (for want of better knowledge) that predictors are statistically independent, and estimate combination probabilities from marginal probabilities (using the reshape long command in Stata). We then input the .ster file and the practice scenario datasets, and used the *margprev* add-on package, and the *parmby* module of the *parmest* add-on package, to estimate marginal prevalence of disease, using the model to make out-of-sample predictions for each practice to produce the synthetic estimates for that package. An example of this can be found in Newson (2014). [58] Confidence intervals (CIs) are also calculated for each local estimate by the packages.

## Validation of local estimates

### Internal validation of local estimates

The local estimates can also be validated by aggregating them to the lowest geography available in the raw data and comparing them. We also present these results. We have over time increased the number of variables used in the local models as more local data has become available. However as more variables are added we need to take account of the joint effects of multiple risk factors, i.e. it assumes they operate independently. Estimation of the joint effects of multiple risk factors is complex for several reasons. In particular, some of the effects of more distal risk factors are mediated through intermediate factors.

When estimating the total effects of individual distal factors on disease, both mediated and direct effects should be considered, because in the presence of mediated effects, controlling for the intermediate factor would attenuate the effects of the more distal one. [59] When estimating the joint effects of the more distal factor and the intermediate one, the mediated and direct effects should be separated, especially if the intermediate factor is affected by other distal factors.

Finally, there can be collinearity between exposure to various risk factors, meaning that one can be linearly predicted from the others with a substantial degree of accuracy. In this situation the coefficient estimates of the multiple regression may change erratically in response to small changes in the model or the data. Collinearity does not reduce the predictive power or reliability of the model as a whole, at least within the sample data set; it only affects calculations regarding individual predictors. That is, a multiple regression model with correlated predictors can indicate how well the entire bundle of predictors predicts the outcome variable, but it may not give valid results about any individual predictor, or about which predictors are redundant with respect to others.

### External validation/comparison with QOF prevalence

The funding for the project does not include an in-depth external validation. For example, this could be carried out by obtaining an extract from a similar dataset e.g. applying the HSfE prevalence models’ equations to (possibly) English Longitudinal Study of Ageing (ELSA) data, but ELSA includes only over 50s, and there are few other data sources for depression. Once the NPMS 2014 data is available, this would be a suitable source for external validation.

However another external data source is the Quality & Outcomes Framework (QOF) GP-diagnosed hypertension prevalence, as the local population estimates will most frequently be compared to QOF-registered prevalence. We carried out an external validation by examining the practice level depression prevalence estimates and QOF registered prevalence. The former can obviously be compared with diagnosed depression prevalence from the model, taking into account that the HSfE definition was derived from the number of patients that reported being told by a nurse or doctor that they had had depression.

We therefore carried out a disagreement analysis between model-estimated and QOF prevalence (%) of diagnosed hypertension in CCGs and practices. We estimated three principal components of disagreement (discordance as measured by Kendall's tau-a, bias as measured by percentile differences, and scale discrepancy as measured by Kendall’s tau-a between the mean of the two prevalences and the difference between the two prevalences). Confidence intervals were calculated for Kendall’s tau-a using the methods of Newson (2006a) [92], and for percentile differences using the methods of Newson (2006b) [93].

Finally, it is sometimes possible to compare Regional breakdowns as a form of external validation. However this is very difficult with HSfE and QOF data because NHS Regions and old Government Office Regions, which are used in HSfE) are no longer co-terminous.

# Results

## Population and baseline characteristics

Figure 1: Depression population flowchart (Health Survey from England 2014)

10,080 participants

8077 participants

5,485 participants

2003 participants under 16 years old

2596 participants did not answer the nurse questionnaire regarding depression

The flowchart in Figure 1 above describes all the steps made in order to obtain our final population. From the HSfE 2014, we obtained a total of 10,080 respondents. From these, observations from individuals below the age of 16 (2003 respondents) were dropped from the analysis. Additionally, all participants who did not answer the questionnaire regarding depression (2596 respondents) were also dropped. Our analysis was therefore performed in a sample of 5,485 participants.

## Statistical analysis

### Descriptive Statistics for the final population

Table 6 show the descriptive statistics for demographics and main risk factors in our final population, stratified for the 2014 and in total. In order to analyse the data from HSfE amidst missing data, multiple imputation was performed on the final population in the 2014 survey to replace the missing data. The final population after carrying out 20 multiple imputations was 115,185 i.e. there were 20 copies of the original dataset. The prevalence of “reported experiencing depression including post-natal depression, but were not diagnosed by a health professional” (DiagCX7) was 27.49%, which includes both with or without self-report diagnosis. The actual question is: *“Do you think that you have ever experienced any of these? Depression*. The prevalence of self- (patient)-reported doctor-diagnosed depression (BCond\_7\_HProf) was 18.80%. The actual question is: *“Did a doctor, psychiatrist or other professional tell you that you had the following condition: Depression?*” There is also a question: *“In the last 12 months have you had: Depression?”*

We decided that, although the first question could be used to measure the extent of undiagnosed depression, use of the second question i.e. health professional diagnosed depression was more consistent with the outcome definitions for other models and best for gauging the effectiveness of case-finding by practices. Another option might be to sum or in other ways combine the two definitions to get the best estimate of overall population prevalence at any point in time.

Table : Distribution of risk and protective factor among the final population (after performing multiple imputation), HSfE 2014

| Risk/protective factor | | N | Percentage |
| --- | --- | --- | --- |
| Outcome measures | |  |  |
| Reported experiencing depression including post-natal depression (DiagCX7, but were not diagnosed by a health professional | |  |  |
|  | Yes, with self-report diagnosis | 1,107 | 20.18 |
|  | Yes, without self-report diagnosis | 401 | 7.31 |
|  | No | 3,977 | 72.51 |
| Self- (patient)-reported doctor-diagnosed depression (BCond\_7\_HProf) | |  |  |
|  | Yes | 1,031 | 18.80 |
|  | No | 4,454 | 81.20 |
| Risk/protective factors |  |  |  |
| Sex |  |  |  |
|  | Male | 2,435 | 44.39 |
|  | Female | 3,050 | 55.61 |
| Age |  |  |  |
|  | 16-24 | 432 | 7.88 |
|  | 25-34 | 715 | 13.04 |
|  | 35-44 | 911 | 16.61 |
|  | 45-54 | 1,019 | 18.58 |
|  | 55-64 | 880 | 16.04 |
|  | 65-74 | 913 | 16.65 |
|  | 75+ | 615 | 11.21 |
| Top qualification | |  |  |
|  | NVQ4/NVQ5/Degree or equiv | 1,411 | 25.75 |
|  | Higher ed below degree | 655 | 11.95 |
|  | NVQ3/GCE A Level equiv | 823 | 15.02 |
|  | NVQ2/GCE O Level equiv | 1,081 | 19.73 |
|  | NVQ1/CSE other grade equiv | 209 | 3.81 |
|  | Foreign/other | 71 | 1.30 |
|  | No qualification | 1,229 | 22.43 |
| Employment category | |  |  |
|  | Higher managerial and professional occ | 668 | 12.34 |
|  | Lower managerial and professional occup | 1,236 | 22.83 |
|  | Intermediate occupations | 819 | 15.13 |
|  | Small employers and own account workers | 477 | 8.81 |
|  | Lower supervisory and technical occupat | 361 | 6.67 |
|  | Semi-routine occupations | 932 | 17.22 |
|  | Routine occupations | 689 | 12.73 |
|  | Never worked and long term unemployed | 94 | 1.74 |
|  | Other | 137 | 2.53 |
| Ethnicity | |  |  |
|  | White | 4,958 | 90.39 |
|  | Mixed | 56 | 1.02 |
|  | South Asian | 229 | 4.18 |
|  | African/Afro-Caribbean | 118 | 2.15 |
|  | Other background |  |  |
| IMD |  |  |  |
|  | 0.53->8.49 [least deprived] | 1,259 | 22.95 |
|  | 8.49->13.79 | 1,105 | 20.15 |
|  | 13.79->21.35 | 1,118 | 20.38 |
|  | 21.35->34.17 | 1,069 | 19.49 |
|  | 34.17->87.80 [most deprived] | 934 | 17.03 |
|  | 0.53->8.49 [least deprived] | 1,259 | 22.95 |
| BMI |  |  |  |
|  | Under 18.5 | 103 | 1.98 |
|  | 18.5 and below 25 | 2,150 | 41.39 |
|  | 25 and below 30 | 1,908 | 36.73 |
|  | 30 and below 40 | 930 | 17.91 |
|  | Over 40 | 103 | 1.98 |
| Marital Status | |  |  |
|  | Single | 1,953 | 35.62 |
|  | Married or cohabitees | 3,530 | 64.38 |
| Self-reported alcohol consumption | |  |  |
|  | Less than once a week | 2,626 | 48.05 |
|  | Once or twice a week | 1,417 | 25.93 |
|  | Three or four times per week | 694 | 12.70 |
|  | Five or more times per week | 728 | 13.32 |
| Limiting long-lasting illness | |  |  |
|  | No limiting long-lasting illness | 4,041 | 73.75 |
|  | Limiting long-lasting illness | 1,438 | 26.25 |
| Physical activity (Tertile of moderate or vigorous intensive minutes of activity per week) | |  |  |
|  | Low | 1,723 | 36.51 |
|  | Medium | 1,563 | 33.12 |
|  | High | 1,433 | 30.37 |
| Reported ever having experienced alcohol or drug dependence | |  |  |
|  | Yes | 128 | 2.33 |
|  | No | 5,357 | 97.67 |

### Missing values

Table 7 shows the number and percentage of missing values in total and stratified per year. The biggest percentage of missing values was found in the variable self-reported physical activity, followed by for self-reported BMI. With these two exceptions, all other variables missed less than 1.5% of their observations. Missing values from the variable origin where recoded to be considered within the category “white”, while physical activity and smoking status were finally excluded from both models. The remaining missing values where inferred for analysis using multiple imputation.

Table 7: Missing values in the final population.

| Risk or Protective factor | 2014 | |
| --- | --- | --- |
|  | Frequency | %\* |
| Top qualification | 6 | 0.11 |
| Employment category | 72 | 1.31 |
| BMI | 291 | 5.31 |
| Alcohol consumption | 20 | 0.36 |
| Limiting long-lasting illness | 6 | 0.11 |
| Physical activity | 766 | 13.97 |

\*Weighted

### Univariable analyses

Table 8 shows the results of univariate models for individual risk and protective factors, and the outcome of diagnosed depression.

Table : Univariate analysis of risk factors for diagnosed depression (2014)

| **Predictor variable** | Odds ratio | P>z | [95% Conf. Interval] |
| --- | --- | --- | --- |
| Sex |  |  |  |
| Female | 1.91 | 0.000 | 1.63-2.23 |
| Age |  |  |  |
| 25-34 | 1.42 | 0.052 | 1.00-2.02 |
| 35-44 | 2.01 | 0.000 | 1.46-2.77 |
| 45-54 | 2.19 | 0.000 | 1.59-3.03 |
| 55-64 | 2.61 | 0.000 | 1.86-3.68 |
| 65-74 | 1.84 | 0.001 | 1.31-2.59 |
| 75+ | 0.98 | 0.913 | 0.66-1.45 |
| Top qualification |  |  |  |
| Higher education below degree | 0.88 | 0.383 | 0.66-1.17 |
| NVQ3/GCE A Level equiv. | 1.27 | 0.034 | 1.02-1.59 |
| NVQ2/GCE O Level equiv. | 1.43 | 0.001 | 1.15-1.79 |
| NVQ1/CSE other grade | 1.38 | 0.101 | 0.94-2.01 |
| Foreign/other | 1.20 | 0.587 | 0.62-2.30 |
| No qualification | 1.30 | 0.020 | 1.04-1.62 |
| Employment category |  |  |  |
| Lower managerial and professional | 1.21 | 0.184 | 0.91-1.59 |
| Intermediate occupations | 1.31 | 0.096 | 0.95-1.80 |
| Small employers and own account | 1.16 | 0.419 | 0.81-1.66 |
| Lower supervisory and technical | 1.36 | 0.114 | 0.93-1.98 |
| Semi-routine occupations | 1.71 | 0.000 | 1.27-2.29 |
| Routine occupations | 1.67 | 0.001 | 1.22-2.27 |
| Never worked and long-term unemployed | 0.41 | 0.037 | 0.18-0.95 |
| Other | 0.22 | 0.003 | 0.08-0.60 |
| Ethnicity |  |  |  |
| Mixed | 0.19 | 0.003 | 0.06-0.57 |
| South Asian | 0.40 | 0.000 | 0.25-0.63 |
| African/Afro-Caribbean | 0.22 | 0.000 | 0.11-0.46 |
| Other background | 0.24 | 0.000 | 0.12-0.48 |
| IMD |  |  |  |
| 8.49->13.79 | 1.08 | 0.533 | 0.84-1.40 |
| 13.79->21.35 | 1.21 | 0.158 | 0.93-1.57 |
| 21.35->34.17 | 1.2 | 0.162 | 0.93-1.57 |
| 34.17->87.80 | 1.51 | 0.001 | 1.17-1.94 |
| BMI |  |  |  |
| 18.5 and below 25 | 1.08 | 0.820 | 0.58-2.04 |
| 25 and below 30 | 1.29 | 0.417 | 0.69-2.41 |
| 30 and below 40 | 2.16 | 0.017 | 1.14-4.05 |
| Over 40 | 3.51 | 0.001 | 1.67-7.36 |
| Marital Status |  |  |  |
| Married or cohabitees | 0.72 | 0.000 | 0.62-0.85 |
| Alcohol consumption (times per week) |  |  |  |
| Once or twice a week | 0.90 | 0.260 | 0.74-1.09 |
| Three or four times per week | 0.90 | 0.396 | 0.71-1.15 |
| Five or more times per week | 0.91 | 0.450 | 0.71-1.16 |
| Limiting long-lasting illness |  |  |  |
| Limiting long-lasting illness | 3.14 | 0.000 | 2.65-3.73 |
| Physical activity (Tertile of moderate or vigorous intensive minutes of activity per week) |  |  |  |
| Medium | 0.81 | 0.032 | 0.67-0.981 |
| High | 0.77 | 0.007 | 0.63-0.93 |
| Alcohol or drug dependence |  |  |  |
| No | 0.09 | 0.000 | 0.05-0.17 |
| Anxiety |  |  |  |
| Self-reported | 9.15 | 0.000 | 7.04-11.89 |
| Yes, not self-reported | 1.67 | 0.012 | 1.12-2.49 |

### Multivariable analyses

Table 9 shows the ORs of fitting the multivariable logistic regression model. Risk factors are female gender, middle (35-64) age groups, limiting long-lasting illness, and self-reported anxiety. Protective factors are age 75+, being from an ethnic minority, never worked, long term unemployed, and other occupational groups, (which may include retired people) being married or cohabiting, and no alcohol or drug dependence. Of course lower rates of diagnosed depression may be a results of cultural factors related to seeking medical care for depression.

Table : Multivariate analysis for diagnosed depression, HSfE 2014, complete final model

| Predictor Variable | Odds Ratio | P>z | 95% Confidence Intervals | |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| Sex |  |  |  |  |
| Female | 2.00 | 0.000 | 1.67 | 2.39 |
| Age |  |  |  |  |
| 25-34 | 1.06 | 0.785 | 0.68 | 1.64 |
| 35-44 | 1.62 | 0.030 | 1.04 | 2.51 |
| 45-54 | 1.65 | 0.022 | 1.07 | 2.54 |
| 55-64 | 1.60 | 0.033 | 1.04 | 2.45 |
| 65-74 | 1.16 | 0.492 | 0.75 | 1.79 |
| 75+ | 0.44 | 0.001 | 0.27 | 0.70 |
| Ethnicity |  |  |  |  |
| Mixed | 0.26 | 0.010 | 0.09 | 0.72 |
| South Asian | 0.52 | 0.013 | 0.31 | 0.87 |
| African/Afro-Caribbean | 0.19 | 0.000 | 0.09 | 0.40 |
| Other background | 0.27 | 0.000 | 0.12 | 0.54 |
| BMI |  |  |  |  |
| 18.5 and below 25 | 0.90 | 0.770 | 0.46 | 1.79 |
| 25 and below 30 | 1.09 | 0.800 | 0.57 | 2.14 |
| 30 and below 40 | 1.52 | 0.230 | 0.77 | 3.00 |
| Over 40 | 2.21 | 0.063 | 0.96 | 5.11 |
| Employment category |  |  |  |  |
| Lower managerial and professional | 0.84 | 0.261 | 0.62 | 1.14 |
| Intermediate occupations | 0.81 | 0.243 | 0.56 | 1.16 |
| Small employers and own account | 0.85 | 0.402 | 0.57 | 1.25 |
| Lower supervisory and technical | 1.04 | 0.866 | 0.68 | 1.58 |
| Semi-routine occupations | 1.10 | 0.588 | 0.78 | 1.56 |
| Routine occupations | 1.09 | 0.655 | 0.76 | 1.56 |
| Never worked and long term unemployed | 0.33 | 0.016 | 0.13 | 0.81 |
| Other | 0.19 | 0.004 | 0.06 | 0.58 |
| Limiting long-lasting illness |  |  |  |  |
| Limiting long-lasting illness | 2.50 | 0.000 | 2.07 | 3.04 |
| Alcohol or Drug dependence |  |  |  |  |
| No | 0.19 | 0.000 | 0.09 | 0.38 |
| Anxiety |  |  |  |  |
| Self-reported | 6.65 | 0.000 | 4.87 | 9.09 |
| Yes, not self-reported | 1.49 | 0.090 | 0.94 | 2.36 |
| Marital status |  |  |  |  |
| Married or cohabitees | 0.67 | 0.000 | 0.56 | 0.81 |
| IMD |  |  |  |  |
| 8.49->13.79 | 1.01 | 0.947 | 0.77 | 1.33 |
| 13.79->21.35 | 0.95 | 0.734 | 0.71 | 1.27 |
| 21.35->34.17 | 1.05 | 0.721 | 0.79 | 1.40 |
| 34.17->87.80 [most deprived] | 1.26 | 0.140 | 0.93 | 1.69 |

### Tests for interactions and collinearity

Interactions between the different variables of interest were tested in our “Model 1”, which included all the variables found significant in the multivariable analysis. Interactions were tested individually using Wald tests to determine whether they were significant or not. Interactions found significant were included in Model 1, creating a set of alternative models for which performance was compared using ROC curves. After testing a possible interaction, Wald tests were also conducted to confirm whether the original variables should be kept or not in the model. This process was performed for both models, the most significant interactions for depression (Table 10) and (**Error! Reference source not found.**) are shown below. In both tables only significant p-values from the interaction between two variables are shown, however, the value of the Wald test refers to the entire interaction variable.

Table 10: tests for interactions/effect modification in multivariable Model 1

| Interaction | | Interaction status/p-value | Wald test (Prob>F) |
| --- | --- | --- | --- |
| Sex | **Age** |  |  |
| Female | 75+ | 0.036 |  |
| Sex | **Anxiety** |  |  |
| Female | Self-reported | 0.036 |  |
| Male | Lower supervisory and technical occupations |  |
| Male | Semi-routine occupations |  |
| Male | Routine occupations |  |
| Sex | **IMD** |  |  |
| Male | 34.17->87.80 [most deprived] |  |  |
| Sex | **BMI** |  |  |
| Male | 18.5 to below 25 |  |  |
| Male | 25 to below 30 |  |
| Male | 30 to below 40 |  |
| Male | Over 40 |  |
| Age | **Limiting long-lasting illness** |  |  |
| 75+ | Limiting long-lasting illness |  |  |
| Ethnicity | **Limiting long-lasting illness** |  |  |
| Mixed | Limiting long-lasting illness |  |  |
| South Asian | Limiting long-lasting illness |  |
| IMD | **Limiting long-lasting illness** |  |  |
| 13.79->21.35 | No limiting long-lasting illness |  |  |
| 21.35->34.17 | Limiting long-lasting illness |  |
| 34.17->87.80 | Limiting long-lasting illness |  |

### Stepwise forward and backward selection model

We have fitted several models in order to obtain the most parsimonious model. The details of the different models are shown in Annex 2.

## Internal validation/discrimination

### ROC curves

We next evaluated model performance using the ***predict*** command and receiver operating characteristics (ROC) curves for the various models. The best ROC curve which predicts data perfectly will touch the top-left corner of the plot (area 1.0), and the larger the area under the ROC curve the better the prediction. An area of 0.5 signifies a prediction no better than chance. However, as noted in the Methods, the choice of variables and hence the final model also depends on the availability of local data, so the final local model will not predict as well as the optimal/”gold standard” model i.e. shows the ROC curves for the final/”gold standard” model and the local model. Area under the curve is fair at 0.707. The ROC curves for other models are shown in Annex 3.

Table : comparison of different models

| Model description | Model | ROC area | SE | 95% CI |
| --- | --- | --- | --- | --- |
| Auto stepwise forward and backward model with deprivation data | M1\* | 0.7562 | 0.0085 | 0.73966-0.77280 |
| Auto stepwise forward and backward model without deprivation data | M2\*\* | 0.7557 | 0.0085 | 0.73911-0.77234 |
| Auto stepwise forward and backward model with deprivation data but without alcohol/drug dependence, anxiety and marital status data | M3\*\*\* | 0.7185 | 0.0089 | 0.70114-0.73587 |

\*Model 1 includes the variables: sex, age, ethnicity, self-reported BMI, occupation, limiting long-lasting illness, alcohol/drug dependence, anxiety, marital status and IMD.

\*\* Model 2 includes the variables: sex, age, ethnicity, self-reported BMI, occupation, limiting long-lasting illness, alcohol/drug dependence, anxiety and marital status.

\*\*\* Model 3 includes the variables: sex, ages, ethnicity, self-reported BMI, occupation, limiting long-lasting illness and IMD.

Figure 2 illustrates the ROC curve for Model 1.

Figure : ROC curve stepwise forward and backward selection model (M1)



Figure 3 illustrates the ROC curve for Model 2.

Figure : ROC curve stepwise forward and backward selection model (M2)



Figure 4 illustrates the ROC curve for Model 3.

Figure : ROC curve automatic stepwise forward and backward selection model (M3)

### Predicted probabilities of having depression

We could use the HSfE automatic stepwise forward and backward models combined to predict the probability of individual having depression in HSfE data set, with box plots showing the predicted probability of people having depression among the non-depression and depression groups. Since we have a binary response model, we can choose a cut-off point on the predicted probability to separate the predicted depression cases (with higher predicted probability) from the predicted non-depression cases (with lower predicted probability).

### Sensitivity and specificity analysis

The sensitivity/specificity versus probability cut-off plot shows us the corresponding sensitivity and specificity in each possible probability cut-off point. Higher sensitivity would usually yield low specificity and vice versa, the rule of thumb is to choose a cut-off probability to maximize both. We choose the cut-off probability where sensitivity and specificity lines cross.

### Collinearity

Variance inflation factor (VIF) is used to check for collinearity between independent variables within the regression model. If the VIF value is greater than 10 it suggests collinearity is present. Table 10 shows the VIF values among the risk factor variables within the hypertension model. There is no evidence of collinearity in the model.

Table : VIF values

| Variable | VIF | SQRT VIF | Tolerance | R-squared |
| --- | --- | --- | --- | --- |
| Sex | 1.07 | 1.03 | 0.94 | 0.06 |
| Age | 1.47 | 1.21 | 0.68 | 0.32 |
| Education level | 1.31 | 1.15 | 0.76 | 0.24 |
| Occupation | 1.20 | 1.10 | 0.83 | 0.17 |
| Ethnicity | 1.09 | 1.04 | 0.92 | 0.08 |
| IMD | 1.14 | 1.07 | 0.88 | 0.12 |
| BMI | 1.06 | 1.03 | 0.95 | 0.06 |
| Alcohol consumption | 1.15 | 1.07 | 0.87 | 0.13 |
| Reported experiencing depression | 5.77 | 2.40 | 0.17 | 0.83 |
| Alcohol/Drug dependence | 1.05 | 1.02 | 0.96 | 0.04 |
| Physical activity | 1.11 | 1.05 | 0.90 | 0.10 |
| Self-reported diagnosed depression | 5.70 | 2.39 | 0.18 | 0.82 |
| Limiting long-lasting illness | 1.23 | 1.11 | 0.81 | 0.19 |
| Marital status | 1.11 | 1.05 | 0.90 | 0.10 |
| Anxiety | 1.06 | 1.03 | 0.94 | 0.06 |
| Mean | 1.77 |  |  |  |

## Local prevalence estimates

Local estimates and CIs were calculated by the inverse probability weights method.[60] Table 13 shows descriptive statistics for the practice local estimates. These showed a mean prevalence of 15.78%.

Table : summary statistics for practice estimates

|  |  |
| --- | --- |
| Percentile | Smallest/Largest |
| 1% | 8.00 |
| 5% | 9.64 |
| 10% | 10.85 |
| 25% | 13.68 |
| 50% | 15.78 |
| 75% | 17.35 |
| 90% | 18.78 |
| 95% | 19.65 |
| 99% | 20.90 |

### External validation of local estimates

As noted previously it is difficult to compare prevalence derived from the raw HSfE dataset Regions with corresponding QOF register data for England NHS Regions. However we have compared practice-level estimates of depression prevalence with the corresponding QOF register. The current QOF depression indicator is DEP003, the percentage of patients aged 18 or over with a new diagnosis of depression in the preceding 1 April to 31 March, who have been reviewed not earlier than 10 days after and not later than 56 days after the date of diagnosis. The actual disease register for the depression indicator for the purpose of calculating the register indicator, known as the Adjusted Practice Disease Factor (APDF), is defined as all patients aged 18 or over, diagnosed on or after 1 April 2006, who have an unresolved record of depression in their patient record. Many patients who have had depression will have the episode recorded as resolved and will therefore not appear in the denominator. The average percentage difference between the local estimates, which use a definition of ever having depression, and QOF registers is 8.16%. In general the local estimates are higher than the registered prevalence, as we would expect given the model we developed.

The Kendall's tau-a between model-estimated and QOF prevalence of diagnosed depression for 75520 practices was 0.260 (95% CIs 0.246-0.275)1.1x10-232, Therefore, in our sample of practices, a random pair of practices was 26% percent more likely to be concordant (the practice with the higher QOF prevalence also having the higher model-estimated prevalence) than to be discordant (the practice with the higher QOF prevalence having the lower model-estimated prevalence). **Error! Reference source not found.** shows percentile differences between model-estimated and QOF prevalence of diagnosed hypertension. Note that all percentile differences are positive, even percentile zero, implying that every model-based prevalence was greater than the corresponding QOF prevalence for the same CCG.

Table : percentile differences between model-estimated and QOF prevalence of diagnosed hypertension

|  |  |  |
| --- | --- | --- |
| Percent | Percentile | (95% CIs) |
| 0 | -28.7 | -28.7) |
| 25 | 5.7 | (5.7-5.8) |
| 50 | 8.1 | (8.0-8.2 |
| 75 | 10.4 | (10.3-10.5) |
| 100 | 20.2 | (20.2-20.2) |

The best way to display the data is to plot the difference between the measurements by the two methods for each subject against their mean, creating a Bland-Altman plot. This plot for the local patient-reported doctor-diagnosed prevalence (**Error! Reference source not found.**) shows explicitly the lack of agreement. The difference between the estimates is close to 8%. This is plausible if (a) a significant proportion of diagnosed depression remits, so patients are taken off registers and/or (b) some patients with hypertension have a diagnosis made, but then change GP, or the current GP has told the patient they have depression but does not record it on the depression register (even they may even be treated with antidpressants). **Error! Reference source not found.** shows a Bland-Altman plot for model-estimated and QOF prevalence of diagnosed hypertension.

Figure 5 is a scatter plot of model-estimated and QOF prevalence of diagnosed hypertension. The Kendall tau-a between the mean of the two prevalence rates and the difference between the model-estimated and QOF prevalence rates in 7552 practices was -0.091 (95% CIs-0.091-0.059) means and model-QOF differences are 9 percent more likely to be concordant than to be discordant. This indicates that model-estimated prevalence rates have a higher variability than QOF prevalence.

Figure : scatter plot of model-estimated and QOF prevalence of diagnosed hypertension

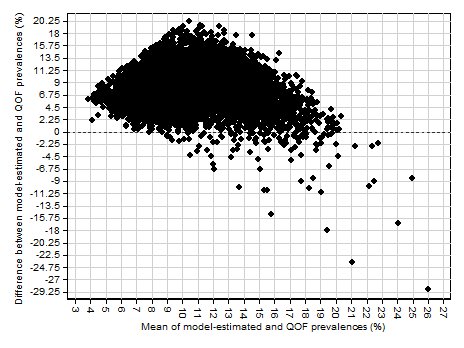
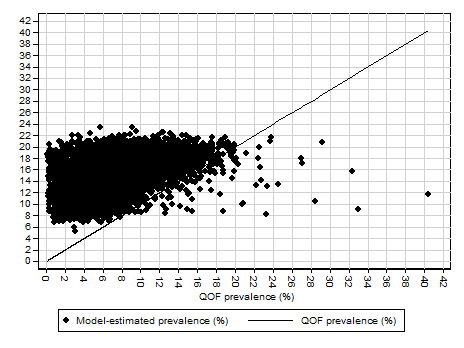
Figure 6 shows a Bland-Altman plot for CCG-level model-estimated and QOF prevalence of diagnosed hypertension. The Kendall tau-a between the mean of the two prevalences and the difference between the model-estimated and QOF prevalences was -0.075 (95% CI -0.091 to -0.059; p= 3.0 x 10-19), showing that prevalence means and model-QOF differences are 7.5 percent more likely to be concordant than to be discordant. This indicates that model-estimated prevalences have a higher variability than QOF prevalences, as suggested by the upwards trend in the scatter plot.

Figure : scatter plot of practice -level model-estimated and QOF prevalence of diagnosed depression at practice levels.



# Discussion

## Summary

There were two options for the depression outcome measure to be used. The first was prevalence of *“reported experiencing depression including post-natal depression, but were not diagnosed by a health professional”* (DiagCX7, prevalence 27.49%), which includes both with or without self-report diagnosis. The actual question is: *“Do you think that you have ever experienced any of these? Depression*. The prevalence of self- (patient)-reported doctor-diagnosed depression (BCond\_7\_HProf) was 18.80%. The actual question is: *“Did a doctor, psychiatrist or other professional tell you that you had the following condition: Depression?”* There is also a question: *“In the last 12 months have you had: Depression?”*

We decided that, although the first question could be used to measure the extent of undiagnosed depression, use of the second question i.e. health professional diagnosed depression was more consistent with the outcome definitions for other models and best for gauging the effectiveness of case-finding by practices. Depending on the objectives of the model, another future option might be to sum or in other ways combine the two definitions to get the best estimate of overall population prevalence at any point in time. The question: *“In the last 12 months have you had: Depression?”* could also be used to estimate local incidence or point prevalence. This could be more useful in the case of an episodic illness.

After conducting a literature review to inform the analysis, in order to generate local estimates for the prevalence of depression in England several models were fitted using data from HSfE 2014. All the known risk or protective factors available in HSfE were tested and either included or excluded from our models using Wald tests and likelihood ratios.

Basing our decision in the performance of each model and the variables available at the local level, a “complete model” (the model with the best possible performance) and a “local model” (the model with the best performance only using variables locally available) were fitted for diagnosed depression. The information from the “local model” and local data retrieved from different sources were used with the inverse probability weights method in order to obtain local prevalence estimates. The average of practice-level prevalence estimates is 15.33%, which is slightly below that in the raw HSfE dataset. This can occur if there are differences in risk factor levels between the HSfE and English populations.

The episodic nature of depression is also a consideration, certainly in terms of comparing our estimates with QOF prevalence, as the latter is meant to exclude resolved depression, but the HSfE variable we used is basically lifetime prevalence. As there is no QOF or other national data on resolved depression it is difficult to compare the two, and we certainly expected the local estimates to be considerably higher, which they were. In this situation annual incidence or point prevalence may also be useful and it would be possible to model it with the 12 month HSfE question.

## Strengths and limitations

### Strengths

In terms of the national data source, strengths of this analysis include the use of a nationally-representative health survey, and rigorous measurement methods by trained survey nurses. HSfE also includes a broad range of risk factor data.

### Limitations

Several possible limitations should be acknowledged when considering the performance of this model, these are mostly in relation to the source of the data used to create the model (HSfE) and to generate the local estimates (Census data from 2011 and others).

In the first place, as with most household surveys, nonresponse may introduce volunteer/respondent bias. Furthermore, individuals living in nursing homes, prisons or military camps, who can display clearly different demographics, were not included in this survey. This is likely to have at the local level for some areas. In addition, the conditions in which the blood pressure measurements were performed might affect results obtained through variations such as the hour of the day in which the measurements took place, or situations such as respondents suffering from white-coat depression. Moreover, it is possible that other significant variables have been excluded from the complete models, due to the characteristics of the HSfE data.

As observed in our model comparison, the performance of the local models is inferior to the complete models. This is a result of some of the limitations of using local data, as not all the variables in the complete model where available locally in a valid format, or changes (e.g. in the differences between categories) were required in order to adapt our model to the available data. More importantly perhaps, interactions, which were observed to improve the performance of the complete models to some extent, could not be included in the final local models.

Another limitation arises from the use of local data, retrieved from public sources (like the 2011 Census) that do not provide annual updates. In addition to evident changes in population demographics happening over time, each year new postcodes are added, and the relation between LSOAs, MSOAs and GP practices potentially changes. This changes produce variations a mismatches when trying to link LSOA/MSOA information to practices.

Finally, if particular cultural groups are less likely to seek medical care for depression (for example ethnic minorities) this will of course lead to their fewer reports of doctor-diagnosed disease.

# References

1. Thomas L, Kessler D, Campbell J, Morrison J, Peters TJ, Williams C, Lewis G, Wiles N. Prevalence of treatment-resistant depression in primary care: cross-sectional data. British Journal of General Practice 2013;63(617):e852-e58. Link: <http://bjgp.org/content/bjgp/63/617/e852.full.pdf>

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3839394/pdf/bjgpdec2013-63-617-e852.pdf>.

2. Veerman JL, Dowrick C, Ayuso-Mateos JL, Dunn G, Barendregt JJ. Population prevalence of depression and mean Beck Depression Inventory score. The British Journal of Psychiatry 2009;195(6):516-19. Link: <http://bjp.rcpsych.org/content/bjprcpsych/195/6/516.full.pdf>.

3. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. The Lancet;370(9590):851-58. Link: <http://www.sciencedirect.com/science/article/pii/S0140673607614159>

<http://ac.els-cdn.com/S0140673607614159/1-s2.0-S0140673607614159-main.pdf?_tid=fc409ff6-2c9f-11e6-ac68-00000aacb362&acdnat=1465297812_5e46c9b8d01e333b177fec566bb901bf>.

4. Davis S, al. E. Annual Report of the Chief Medical Officer 2013, Public Mental Health Priorities: Investing in the Evidence.320. Link: <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/413196/CMO_web_doc.pdf>.

5. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the Price: the cost of mental health care in England to 2026. 2008. Link: <http://www.kingsfund.org.uk/sites/files/kf/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008_0.pdf>.

6. Thomas CM, Morris S. Cost of depression among adults in England in 2000. The British Journal of Psychiatry 2003;183(6):514-19. Link: <http://bjp.rcpsych.org/content/bjprcpsych/183/6/514.full.pdf>.

7. Pickett KE, Wilkinson RG. Inequality: an underacknowledged source of mental illness and distress. The British Journal of Psychiatry 2010;197(6):426-28. Link: <http://bjp.rcpsych.org/content/bjprcpsych/197/6/426.full.pdf>.

8. Bridges S. Annual Report of the Chief Medical Officer 2013. Public Mental Health Priorities: Investing in the Evidence.1:3. Link: <http://www.hscic.gov.uk/catalogue/PUB19295/HSE2014-ch2-mh-prob.pdf>.

9. Ban L, Gibson JE, West J, Fiaschi L, Oates MR, Tata LJ. Impact of socioeconomic deprivation on maternal perinatal mental illnesses presenting to UK general practice. British Journal of General Practice 2012;62(603):e671-e78. Link: <http://bjgp.org/content/bjgp/62/603/e671.full.pdf>

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3459774/pdf/bjgp62-e671.pdf>.

10. Sally Bridges RD, Sara Evans-Lacko, Elizabeth Fuller,, Claire Henderson NI, Natalie Maplethorpe, Jennifer Mindell,, Alison Moody LNF, Keeva Rooney, Nina Sal, Rachel Scantlebury,, Shaun Scholes GT, Raphael Wittenberg. Health Survey for England 2014: Methods & Documentation. 2015. Link: <http://www.hscic.gov.uk/catalogue/PUB19295/HSE2014-Methods-and-docs.pdf>.

11. World Health Organisation. International Classification of Diseases-10 (ICD-10). 2016: World Health Organisation. Link: <http://apps.who.int/classifications/icd10/browse/2016/en>.

12. Maske UE, Buttery AK, Beesdo-Baum K, Riedel-Heller S, Hapke U, Busch MA. Prevalence and correlates of DSM-IV-TR major depressive disorder, self-reported diagnosed depression and current depressive symptoms among adults in Germany. Journal of Affective Disorders 2016;190:167-77. Link: <http://www.sciencedirect.com/science/article/pii/S0165032715303049>

<http://ac.els-cdn.com/S0165032715303049/1-s2.0-S0165032715303049-main.pdf?_tid=ff09fbec-2c9f-11e6-b31d-00000aab0f02&acdnat=1465297817_ef26c2af0c7d54513f38b83c095f82ad>.

13. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 2015;5: American Psychiatric Association. Link: <https://www.appi.org/products/dsm-manual-of-mental-disorders>.

14. Excellence NIfHaC. Depression in adults: recognition and management. 2009. Link.

15. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. The Lancet 2009;374(9690):609-19. Link: <http://dx.doi.org/10.1016/S0140-6736(09)60879-5>.

16. NHS Employers. Quality and outcomes framework. 2016: NHS Employers. Link: <http://www.nhsemployers.org/~/media/Employers/Documents/Primary%20care%20contracts/QOF/2016-17/2016-17%20QOF%20guidance%20documents.pdf>.

17. Kupfer DJ, Frank E, Phillips ML. Major depressive disorder: new clinical, neurobiological, and treatment perspectives. The Lancet;379(9820):1045-55. Link: <http://dx.doi.org/10.1016/S0140-6736(11)60602-8>

<http://ac.els-cdn.com/S0140673611606028/1-s2.0-S0140673611606028-main.pdf?_tid=ab7014de-2e49-11e6-9ca7-00000aab0f01&acdnat=1465480645_6c95ea2fe1460130631a08b1e1e9fc99>.

18. Kessler RC, Angermeyer M, Anthony JC, De Graaf RON, Demyttenaere K, Gasquet I, De Girolamo G, Gluzman S, Gureje OYE, Haro JM, Kawakami N, Karam A, Levinson D, Medina Mora ME, Oakley Browne MA, Posada-Villa J, Stein DJ, Adley Tsang CH, Aguilar-Gaxiola S, Alonso J, Lee S, Heeringa S, Pennell B-E, Berglund P, Gruber MJ, Petukhova M, Chatterji S, ÜStÜN TB, For The Who World Mental Health Survey C. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007;6(3):168-76. Link: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2174588/>

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2174588/pdf/wpa060168.pdf>.

19. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet 2007;370(9590):851-8. Link: <http://ac.els-cdn.com/S0140673607614159/1-s2.0-S0140673607614159-main.pdf?_tid=7555ba04-2e51-11e6-b3df-00000aacb35f&acdnat=1465483987_a231facc3dd10b764cfbae831da82702>.

20. Goldberg D. Are general practitioners unable to diagnose depression? The Lancet;374(9704):1818. Link: <http://dx.doi.org/10.1016/S0140-6736(09)62054-7>

<http://ac.els-cdn.com/S0140673609620547/1-s2.0-S0140673609620547-main.pdf?_tid=3669d9be-2e52-11e6-82d8-00000aacb35f&acdnat=1465484311_f3dd57ac8f04232b41e6ae06b862fc20>.

21. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet 2009;374(9690):609-19. Link: <http://ac.els-cdn.com/S0140673609608795/1-s2.0-S0140673609608795-main.pdf?_tid=cdea5f5c-2e52-11e6-aaaa-00000aacb362&acdnat=1465484565_228dbc43f55caafcee23c139c941f183>.

22. Andrews G, Poulton R, Skoog I. Lifetime risk of depression: restricted to a minority or waiting for most? The British Journal of Psychiatry 2005;187(6):495-96. Link: <http://bjp.rcpsych.org/content/bjprcpsych/187/6/495.full.pdf>.

23. Bauer M, Pfennig A. Epidemiology of bipolar disorders. Epilepsia 2005;46 Suppl 4:8-13. Link.

24. Pini S, de Queiroz V, Pagnin D, Pezawas L, Angst J, Cassano GB, Wittchen HU. Prevalence and burden of bipolar disorders in European countries. Eur Neuropsychopharmacol 2005;15(4):425-34. Link: <http://ac.els-cdn.com/S0924977X0500074X/1-s2.0-S0924977X0500074X-main.pdf?_tid=0cd03422-2eed-11e6-9e0b-00000aab0f27&acdnat=1465550814_0a693f782284649b609365f0acf59680>.

25. !!! INVALID CITATION !!! {Merikangas, 2007 #198;Kupfer, 2005 #203}. Link.

26. Angst J, Gamma A, Benazzi F, Ajdacic V, Eich D, Rossler W. Toward a re-definition of subthreshold bipolarity: epidemiology and proposed criteria for bipolar-II, minor bipolar disorders and hypomania. J Affect Disord 2003;73(1-2):133-46. Link: <http://ac.els-cdn.com/S0165032702003221/1-s2.0-S0165032702003221-main.pdf?_tid=af4ca3fe-2ef0-11e6-8907-00000aacb361&acdnat=1465552375_69ee4d8fb548fd9f5e5eab11f873710c>.

27. Merikangas KR, Akiskal HS, Angst J, et al. LIfetime and 12-month prevalence of bipolar spectrum disorder in the national comorbidity survey replication. Archives of General Psychiatry 2007;64(5):543-52. Link: <http://dx.doi.org/10.1001/archpsyc.64.5.543>

<http://archpsyc.jamanetwork.com/data/Journals/PSYCH/11842/yoa60082_543_552.pdf>.

28. Bener A, Ghuloum S, Abou-Saleh MT. Prevalence, symptom patterns and comorbidity of anxiety and depressive disorders in primary care in Qatar. Social Psychiatry and Psychiatric Epidemiology 2012;47(3):439-46. Link: <http://dx.doi.org/10.1007/s00127-011-0349-9>

[http://download.springer.com/static/pdf/189/art%253A10.1007%252Fs00127-011-0349-9.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-011-0349-9&token2=exp=1465298811~acl=%2Fstatic%2Fpdf%2F189%2Fart%25253A10.1007%25252Fs00127-011-0349-9.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-011-0349-9\*~hmac=7f3f8da87ec8387d3d2a43c0d60fa8528b0b913570e2a48132e1021c994e976d](http://download.springer.com/static/pdf/189/art%253A10.1007%252Fs00127-011-0349-9.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-011-0349-9&token2=exp=1465298811~acl=%2Fstatic%2Fpdf%2F189%2Fart%25253A10.1007%25252Fs00127-011-0349-9.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-011-0349-9*~hmac=7f3f8da87ec8387d3d2a43c0d60fa8528b0b913570e2a48132e1021c994e976d).

29. Busch MA, Maske UE, Ryl L, Schlack R, Hapke U. Prevalence of depressive symptoms and diagnosed depression among adults in Germany. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 2013;56(5):733-39. Link: <http://dx.doi.org/10.1007/s00103-013-1688-3>

[http://download.springer.com/static/pdf/904/art%253A10.1007%252Fs00103-013-1688-3.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00103-013-1688-3&token2=exp=1465298808~acl=%2Fstatic%2Fpdf%2F904%2Fart%25253A10.1007%25252Fs00103-013-1688-3.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00103-013-1688-3\*~hmac=58d06a8641445d1d713efc11e44ef05a1dc912d893c737c0788d8cd629a0f42e](http://download.springer.com/static/pdf/904/art%253A10.1007%252Fs00103-013-1688-3.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00103-013-1688-3&token2=exp=1465298808~acl=%2Fstatic%2Fpdf%2F904%2Fart%25253A10.1007%25252Fs00103-013-1688-3.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00103-013-1688-3*~hmac=58d06a8641445d1d713efc11e44ef05a1dc912d893c737c0788d8cd629a0f42e).

30. Li D, Zhang D-j, Shao J-j, Qi X-d, Tian L. A meta-analysis of the prevalence of depressive symptoms in Chinese older adults. Archives of Gerontology and Geriatrics 2014;58(1):1-9. Link: <http://www.sciencedirect.com/science/article/pii/S0167494313001350>

<http://ac.els-cdn.com/S0167494313001350/1-s2.0-S0167494313001350-main.pdf?_tid=f6b15580-2c9f-11e6-9fee-00000aab0f01&acdnat=1465297803_ddd40ddb533bf092b30af059962a7fa5>.

31. AYUSO-MATEOS JL, VÁZQUEZ-BARQUERO JL, DOWRICK C, LEHTINEN V, DALGARD OS, CASEY P, WILKINSON C, LASA L, PAGE H, DUNN G, WILKINSON G. Depressive disorders in Europe: prevalence figures from the ODIN study. The British Journal of Psychiatry 2001;179(4):308-16. Link: <http://bjp.rcpsych.org/content/bjprcpsych/179/4/308.full.pdf>.

32. Park JH, Kim KW, Kim M-H, Kim MD, Kim B-J, Kim S-K, Kim JL, Moon SW, Bae JN, Woo JI, Ryu S-H, Yoon JC, Lee N-J, Lee DY, Lee DW, Lee SB, Lee JJ, Lee J-Y, Lee C-U, Chang SM, Jhoo JH, Cho MJ. A nationwide survey on the prevalence and risk factors of late life depression in South Korea. Journal of Affective Disorders 2012;138(1–2):34-40. Link: <http://www.sciencedirect.com/science/article/pii/S0165032711008019>

<http://ac.els-cdn.com/S0165032711008019/1-s2.0-S0165032711008019-main.pdf?_tid=41bcc4de-2cbd-11e6-a04c-00000aab0f6c&acdnat=1465310384_22a68186053aee205f80391ff90ba955>.

33. Crimmins EM, Kim JK, Solé-Auró A. Gender differences in health: results from SHARE, ELSA and HRS. The European Journal of Public Health 2011;21(1):81-91. Link: <http://eurpub.oxfordjournals.org/content/eurpub/21/1/81.full.pdf>

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3023013/pdf/ckq022.pdf>.

34. Kessler RC. Epidemiology of women and depression. Journal of Affective Disorders 2003;74(1):5-13. Link: <http://www.sciencedirect.com/science/article/pii/S0165032702004263>

<http://ac.els-cdn.com/S0165032702004263/1-s2.0-S0165032702004263-main.pdf?_tid=7507c7e8-2d6d-11e6-b10f-00000aacb361&acdnat=1465386062_975e92c18472446166bec9cdcd832eba>.

35. Roberts RE, Kaplan GA, Shema SJ, Strawbridge WJ. Does growing old increase the risk for depression? Am J Psychiatry 1997;154(10):1384-90. Link: <http://ajp.psychiatryonline.org/doi/pdf/10.1176/ajp.154.10.1384>.

36. Brown JSL, Casey SJ, Bishop AJ, Prytys M, Whittinger N, Weinman J. How Black African and White British Women Perceive Depression and Help-Seeking: a Pilot Vignette Study. International Journal of Social Psychiatry 2011;57(4):362-74. Link: <http://isp.sagepub.com/content/57/4/362.abstract>

<http://isp.sagepub.com/content/57/4/362.full.pdf>.

37. Alegria M, Molina KM, Chen CN. Neighborhood characteristics and differential risk for depressive and anxiety disorders across racial/ethnic groups in the United States. Depress Anxiety 2014;31(1):27-37. Link: <http://onlinelibrary.wiley.com/store/10.1002/da.22197/asset/da22197.pdf?v=1&t=ip6yp6k4&s=2c719b8a1d4a3dfa7e3adca447779d9c494fefe3>.

38. Atlantis E, Baker M. Obesity effects on depression: systematic review of epidemiological studies. Int J Obes 2008;32(6):881-91. Link: <http://dx.doi.org/10.1038/ijo.2008.54>

<http://www.nature.com/ijo/journal/v32/n6/pdf/ijo200854a.pdf>.

39. Luppino FS, de Wit LM, Bouvy PF, et al. Overweight, obesity, and depression: A systematic review and meta-analysis of longitudinal studies. Archives of General Psychiatry 2010;67(3):220-29. Link: <http://dx.doi.org/10.1001/archgenpsychiatry.2010.2>

<http://archpsyc.jamanetwork.com/data/Journals/PSYCH/5290/yma90003_220_229.pdf>.

40. Bjelland I, Krokstad S, Mykletun A, Dahl AA, Tell GS, Tambs K. Does a higher educational level protect against anxiety and depression? The HUNT study. Social Science & Medicine 2008;66(6):1334-45. Link: <http://www.sciencedirect.com/science/article/pii/S0277953607006776>

<http://ac.els-cdn.com/S0277953607006776/1-s2.0-S0277953607006776-main.pdf?_tid=440a1c84-2d65-11e6-ade5-00000aacb361&acdnat=1465382543_08371f5e992fc4d730dd45d2bf292d15>.

41. Bulloch AG, Williams JV, Lavorato DH, Patten SB. The relationship between major depression and marital disruption is bidirectional. Depression and Anxiety 2009;26(12):1172-77. Link: <http://dx.doi.org/10.1002/da.20618>

<http://onlinelibrary.wiley.com/store/10.1002/da.20618/asset/20618_ftp.pdf?v=1&t=ip6qtrdv&s=4e5ff599182d14b59c95d1989f209a5ed6f60177>.

42. Beesdo K, Bittner A, Pine DS, Stein MB, Hofler M, Lieb R, Wittchen HU. Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. Arch Gen Psychiatry 2007;64(8):903-12. Link: <http://archpsyc.jamanetwork.com/data/Journals/PSYCH/11847/yoa70008_903_912.pdf>.

43. Kessler RC, Birnbaum H, Bromet E, Hwang I, Sampson N, Shahly V. Age differences in major depression: results from the National Comorbidity Survey Replication (NCS-R). Psychological Medicine 2010;40(02):225-37. Link: <http://dx.doi.org/10.1017/S0033291709990213>

<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=7018792&fileId=S0033291709990213>.

44. Brunes A, Augestad LB, Gudmundsdottir SL. Personality, physical activity, and symptoms of anxiety and depression: the HUNT study. Social Psychiatry and Psychiatric Epidemiology 2013;48(5):745-56. Link: <http://dx.doi.org/10.1007/s00127-012-0594-6>

[http://download.springer.com/static/pdf/889/art%253A10.1007%252Fs00127-012-0594-6.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-012-0594-6&token2=exp=1465383567~acl=%2Fstatic%2Fpdf%2F889%2Fart%25253A10.1007%25252Fs00127-012-0594-6.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-012-0594-6\*~hmac=ce567df2dfeac98e27027e0185140f50fb028472105dde551c903ed0834ddb77](http://download.springer.com/static/pdf/889/art%253A10.1007%252Fs00127-012-0594-6.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-012-0594-6&token2=exp=1465383567~acl=%2Fstatic%2Fpdf%2F889%2Fart%25253A10.1007%25252Fs00127-012-0594-6.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-012-0594-6*~hmac=ce567df2dfeac98e27027e0185140f50fb028472105dde551c903ed0834ddb77).

45. Augestad LB, Slettemoen RP, Flanders WD. Physical Activity and Depressive Symptoms Among Norwegian Adults Aged 20–50. Public Health Nursing 2008;25(6):536-45. Link: <http://dx.doi.org/10.1111/j.1525-1446.2008.00740.x>

<http://onlinelibrary.wiley.com/store/10.1111/j.1525-1446.2008.00740.x/asset/j.1525-1446.2008.00740.x.pdf?v=1&t=ip6qtjsh&s=2c44f44e1cb107a525f528dbaf4d5a69517a60eb>.

46. McKercher C, Sanderson K, Schmidt MD, Otahal P, Patton GC, Dwyer T, Venn AJ. Physical activity patterns and risk of depression in young adulthood: a 20-year cohort study since childhood. Soc Psychiatry Psychiatr Epidemiol 2014;49(11):1823-34. Link: [http://download.springer.com/static/pdf/437/art%253A10.1007%252Fs00127-014-0863-7.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-014-0863-7&token2=exp=1465392024~acl=%2Fstatic%2Fpdf%2F437%2Fart%25253A10.1007%25252Fs00127-014-0863-7.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-014-0863-7\*~hmac=bee25b9630b6fb0a7a5d29b56933801ff5a8ed92b49afdd6ab624a8ea0de542c](http://download.springer.com/static/pdf/437/art%253A10.1007%252Fs00127-014-0863-7.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs00127-014-0863-7&token2=exp=1465392024~acl=%2Fstatic%2Fpdf%2F437%2Fart%25253A10.1007%25252Fs00127-014-0863-7.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs00127-014-0863-7*~hmac=bee25b9630b6fb0a7a5d29b56933801ff5a8ed92b49afdd6ab624a8ea0de542c).

47. Assari S, Burgard S, Zivin K. Long-Term Reciprocal Associations Between Depressive Symptoms and Number of Chronic Medical Conditions: Longitudinal Support for Black-White Health Paradox. J Racial Ethn Health Disparities 2015;2(4):589-97. Link: <http://link.springer.com/article/10.1007/s40615-015-0116-9/fulltext.html>

[http://download.springer.com/static/pdf/332/art%253A10.1007%252Fs40615-015-0116-9.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs40615-015-0116-9&token2=exp=1465383559~acl=%2Fstatic%2Fpdf%2F332%2Fart%25253A10.1007%25252Fs40615-015-0116-9.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs40615-015-0116-9\*~hmac=dd80541a9add6b9d569025f66322fe0df8bd9fdbac4b3d2e27eae2e8c3c42d6c](http://download.springer.com/static/pdf/332/art%253A10.1007%252Fs40615-015-0116-9.pdf?originUrl=http%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs40615-015-0116-9&token2=exp=1465383559~acl=%2Fstatic%2Fpdf%2F332%2Fart%25253A10.1007%25252Fs40615-015-0116-9.pdf%3ForiginUrl%3Dhttp%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs40615-015-0116-9*~hmac=dd80541a9add6b9d569025f66322fe0df8bd9fdbac4b3d2e27eae2e8c3c42d6c).

48. Neckelmann D, Mykletun A, Dahl AA. Chronic insomnia as a risk factor for developing anxiety and depression. SLEEP-NEW YORK THEN WESTCHESTER- 2007;30(7):873. Link: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1978360/pdf/aasm.30.7.873.pdf>.

49. Roberts RE, Deleger S, Strawbridge WJ, Kaplan GA. Prospective association between obesity and depression: evidence from the Alameda County Study. Int J Obes Relat Metab Disord 2003;27(4):514-21. Link: <http://www.nature.com/ijo/journal/v27/n4/pdf/0802204a.pdf>.

50. Ali S, Davies MJ, Taub NA, Stone MA, Khunti K. Prevalence of diagnosed depression in South Asian and white European people with type 1 and type 2 diabetes mellitus in a UK secondary care population. Postgraduate Medical Journal 2009;85(1003):238-43. Link: <http://pmj.bmj.com/content/85/1003/238.abstract>

<http://pmj.bmj.com/content/85/1003/238.full.pdf>.

51. Franzen PL, Buysse DJ. Sleep disturbances and depression: risk relationships for subsequent depression and therapeutic implications. Dialogues in Clinical Neuroscience 2008;10(4):473-81. Link: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3108260/>

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3108260/pdf/DialoguesClinNeurosci-10-473.pdf>.

52. He Y, Zaslavsky AM, Landrum MB, Harrington DP, Catalano P. Multiple imputation in a large-scale complex survey: a practical guide. Statistical Methods in Medical Research 2010;19(6):653-70. Link: <http://smm.sagepub.com/content/19/6/653.abstract>.

53. Li P, Stuart EA, Allison DB. Multiple imputation: A flexible tool for handling missing data. JAMA 2015;314(18):1966-67. Link: <http://dx.doi.org/10.1001/jama.2015.15281>.

54. Royston P. Multiple imputation of missing values. Stata Journal 2004;4(3):227-41. Link: <http://www.stata-journal.com/article.html?article=st0067>

<http://www.stata-journal.com/sjpdf.html?article=st0067>.

55. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338(jun29\_1):b2393. Link: <http://www.bmj.com>

<http://www.bmj.com/cgi/content/full/338/jun29_1/b2393>.

56. Kirkwood BR, Sterne JAC. Regression modelling. In: K M, editor. *Medical Statistics*. USA: Blackwell Publishing company, 2003:339-42.

57. Newson RB. Attributable and unattributable risks and fractions and other scenario comparisons. Stata Journal 2013;13(4):672-98. Link: <http://www.stata-journal.com/article.html?article=st0314>.

58. Newson RB. Scenario comparisons: How much good can we do?, 2014.

59. Ezzati M, Vander Hoorn S, Rodgers A, Lopez AD, Mathers CD, Murray CJL. Estimates of global and regional potential health gains from reducing muliple major risk factors. The Lancet 2003;362(9380):271-80. Link: <http://www.sciencedirect.com/science/article/pii/S0140673603139682>.

60. Mansournia MA, Altman DG. Inverse probability weighting. BMJ 2016;352. Link: <http://www.bmj.com/bmj/352/bmj.i189.full.pdf>.

61. Newson R. Confidence intervals for rank statistics: Somers' *D* and extensions. Stata J 2006;6(3):309-34. Link: <Go to ISI>://WOS:000240514400002.

62. Newson R. Confidence intervals for rank statistics: Percentile slopes, differences, and ratios. Stata J 2006;6(4):497-520. Link: <Go to ISI>://WOS:000242923600004.

# Annex 1: synthetic estimation using Stata

## Synthetic estimates

The description below uses data from ELSA as an example, but the method is the same for any given data source, and was used for the HSfE estimates. The proportion of our population according to age and sex are known. The proportion by educational status, for example, can be applied to these numbers, taking account of the fact that the distribution by educational status differs by age group. This gives estimated proportion by age, sex and educational status. This information is reflected in the variables below (variable names starting m\_noed\_, m\_othed, m\_nqv\_, f\_noed, f\_othed, f\_nqv).

Within Stata, a new set of variables is created, one for each combination of these risk factors pertinent to the logistic regression model for the chosen disease. For instance, if there are two binary variable for age group included in the regression model, then there are three relevant age groups (those the first variable=1, those with the second variable=1, and those where both variables=0 – it is not possible to have both variables =1 since this would imply being in two separate age groups at the same time). With a binary variable for gender included, we would need groups for each gender – but some models don’t include gender, like the one here we are using in this illustration. With one binary variable for educational status included in our predictive model, there are 2 categories for education (those with and those without this specified educational status, which here is no qualifications). The total number of combinations of age/ sex/ education groups then becomes 3x1x2=6. Corresponding to these 6 categories we have 6 variables as follows, which are created by summing sub categories (the categories that have equivalent risk within the model in question):

gen agegp\_23\_edu7\_1\_0=m\_noed\_4564+ f\_noed\_4564

gen agegp\_23\_edu7\_1\_1=m\_noed\_6574+ f\_noed\_6574

gen agegp\_23\_edu7\_1\_2=m\_noed\_75p+ f\_noed\_75p

gen agegp\_23\_edu7\_1\_3=f\_othed\_4564 +f\_nvq\_4564+ m\_othed\_4564+ m\_nvq\_4564

gen agegp\_23\_edu7\_1\_4=f\_othed\_6574 +f\_nvq\_6574+ m\_othed\_6574+ m\_nvq\_6574

gen agegp\_23\_edu7\_1\_5=f\_othed\_75p +f\_nvq\_75p+ m\_othed\_75p+ m\_nvq\_75p

These are calculated based on 3 initial education groups – those with no education ( \_noed\_ variables), those with NQVs only ( \_nvq\_ variables) and those with other education ( \_othed\_ ) although in this model there is no distinction between those with NVQs and other education (since only the binary variable for no education is included in the model). There is also no distinction between males (variables names starting m\_) and females (starting f\_). There is distinction between each of the three age groups (45-64, 65-74 and 75 plus), since both binary variables for age categories are included in this model.

Of course, they could be calculated in any way convenient, provided the result is the anticipated proportion in each age/ sex/ educational group, pertinent to the model in hand. They can be named in any convenient way, providing each has the same name apart from having a different number at the end. This allows use of the reshape command in Stata.

In practice, we do not want to find a synthetic estimate on just one population, but rather on many populations, for instance on each local authority separately. We have a data set containing information on the risk factors in all the different local authorities (LAs) and also other regions, with one line of data per region. The above variables give the proportions for each specified combination of age/ sex/ education categories. There are other variables giving the proportions by each additional risk factor separately (e.g. the proportion of non-smokers, current smokers and ex-smokers).



A reshape long command on the set of 6 agegp\_23\_edu7\_1\_ variables (as defined above) is used as follows:

reshape long agegp\_23\_edu7\_1\_, i(ccg\_code) j(agegp\_23\_edu7\_1)

this gives 6 lines of data per region (since in this example there are 6 categories of age/ sex/ educational status, and 6 corresponding variables) from the starting place of one line of data per region. As well as a variable defining the categories (agegp\_23\_edu7\_1 as named in the j() part of the above command), we now have a variable giving the proportion in each row of data (called variable agegp\_23\_edu7\_1\_ note that this name ends in \_). These proportions were originally 6 variables on each line, and now we have 6 separate lines for each region. (If you look at the data listing above, the row of proportions turns into a column of 6 proportions, then the second row becomes a column of another 6 proportions below the first six, against the second LA code). The i() part of the command gives a unique identifier for each line of data.



For a risk factor, such as smoking status, where the number by age, sex and other risk factors is not known, the proportion of smokers and of ex-smokers in the population is applied to each age/ sex/ educational status group. Another such variable is physical activity (PA) level (palevel), which is in 4 categories, so has 3 corresponding binary variables, all of which are included in this predictive logistic regression model. This is the next one dealt with in practice.

Four relevant variables are created as follows for PA level, with the requirement that they all have the same name, except for the different numbers at the end, as follows:

gen palevelf\_0=1-pa\_low-pa\_mod - pa\_high

gen palevelf\_1=pa\_low

gen palevelf\_2=pa\_mod

gen palevelf\_3=pa\_high

(derived from pa\_ variables for low, moderate and high physical activity levels).

With those 4 variables, a further reshape long command can be applied. Note that we already have 6 lines of data per region. This gives 4 lines of data (one for each PA level) from each line, which gives 6x4=24 lines of data per region now. The i() part of the command that gives the unique identifier now needs to include the age/sex/ education categories variable (agegp\_23\_edu7\_1) as well as the region coding variable (ccg\_code). The j() part tells Stata to name the newly created categorical variable palevelf, which represents the different PA level categories. The palevelf\_ variable (note \_ at end of this name) gives the proportion within each PA level category (these add to one for each la\_code/ agegp\_23\_edu7\_1 combination, i.e. for each set of 4 lines – again the 4 values that are listed horizontally above are now listed vertically into this palevelf\_ column).

reshape long palevelf\_, i(ccg\_code agegp\_23\_edu7\_1) j(palevelf)



Similarly for other risk factors. For this model, the other risk factors are BMI (obese, overweight and not overweight categories), smoking (where only ex-smoking is relevant, smokers and non-smokers are combined), gym membership and socio-economic status (with 3 relevant binary variables, giving 4 categories). Therefore for this model, there are

6 x 4 x 3 x 2 x 2 x 4 = 1152 different combinations of predictor variables. With 6 different “reshape long” commands in total, we end up with 1152 lines of data per region.

The weights for each region can be obtained by multiplying the relevant proportions together.

Weight = (proportion in specified age/ sex/ education category ) x (proportion by PA level) x (proportion by BMI group) x (proportion by smoking status) x (proportion by gym membership) x (proportion by relevant socio-economic status group).

gen xyz= agegp\_23\_edu7\_1\_ \* palevelf\_\* bmicatf2\_\* smokef2\_\* hobby1\_\* ssec8\_

These weights (“xyz”) will sum to one for each region. It is a good idea to check that they do so in practice.

For practical purposes, so that we can use Stata efficiently, it is also necessary to create all the binary variables used in the original logistic regression modelling, used to derive our preferred predictive model with regression coefficients. The names and coding of these variables must be identical to those used in the original data set.

The most complex is recreating age, education and sex variables, since they are combined above for the purposes of the reshape command. For the model in our example, we do not need a sex variable, but we do need the following variables (check with the above commands which define them initially to make sure the appropriate codings are used – the tab2 command below also allows for some checking):

gen agegp2=agegp\_23\_edu7\_1==1 | agegp\_23\_edu7\_1==4

gen agegp3=agegp\_23\_edu7\_1==2 | agegp\_23\_edu7\_1==5

gen educ7=agegp\_23\_edu7\_1==0 | agegp\_23\_edu7\_1==1 | agegp\_23\_edu7\_1==2

tab2 agegp\_23\_edu7\_1 agegp2 agegp3 educ7, missing

For other variables, such as PA level, it is straight forward to create the required binary variables (the tab2 command again allows for some checking):

gen palevelf1=palevelf==1

gen palevelf2=palevelf==2

gen palevelf3=palevelf==3

tab2 palevelf palevelf1 palevelf2 palevelf3, missing

Note on creation of above variables: the right hand side are expressions, such as palevelf==1 – the variable is coded as =1 when this is true and =0 when this is false and including for missing values of palevelf (here we excluded any data with missing values earlier so this does not apply).

With our dataset set up in this way, we can now use Stata’s “predict” command to give us the predicted log odds. For this to work, the last regression that we have undertaken in Stata must be the definitive predictive logistic regression equation for the chosen disease, which requires the dataset used to derive that to be in Stata’s memory at the time. When we use the “predict” command we need the dataset described above (after all the above described transformations), to be in Stata’s memory, since that gives the characteristics of the regions on which we want the synthetic estimates. It would also be possible to programme in the linear equation from the logistic regression manually, but I have not done that, since there is then more scope for errors.

The predict commands gives predicted log odds, and we then find prevalences as follows: exp(log odds) / [1+ exp(log odds)], Exactly as in the Excel VBA code visible in the workbook. Then we find the weighted average of these, averaged across all possible combinations of risk factors, using the weights calculated as above (stored in variable named xyz). Using Stata, the weighted average can be found using the “collapse” command as follows, which results in one line of data per region (using a region identifier as the by() variable).

predict pred\_values, xb

gen pred\_OR=exp(pred\_values)

gen pred\_prev=pred\_OR/(1+pred\_OR)

gen wt\_pred\_prev=pred\_prev\*xyz

collapse (sum) wt\_pred\_prev xyz, by(ccg\_code)

Thus the region is a data set with one line of data per region, with an estimate of prevalence against each region, based on the definitive logistic regression equation.

## Calculating confidence intervals for prevalence estimates using boot strap procedures

There is uncertainty in these synthetic estimates of prevalence based on the imprecision in the estimated coefficients from the logistic regression equations. A boot strap procedure can be used to construct confidence intervals on these synthetic estimates of prevalence, based on the imprecision in these logistic regression coefficients.

**Boot-strap procedures**

The philosophy underlying the boot-strap procedure is to consider that the people included in the data set used to derive the logistic regression equation represent the whole population of possible people. However, the whole population is effectively considered to contain thousands of copies of each of these people.

Boot strap samples are taken from our initial populations (the subsets of the ELSA population that has complete data on appropriate risk factors). The first person to be included in our new boot strap data set is chosen at random from our starting (ELSA) dataset, with each person being equally likely to be chosen. Then the second person to be included in this boot strap data set is chosen at random in the same way, again with each person being equally likely to be chosen. It is noteworthy that the second person to be chosen could be the same person as the person selected first (with probability 1/n where n=sample size, the same probability that any individual will be selected). We then select a 3rd person for our boot strap sample, then a 4th, 5th, 6th, and so on up to an nth person (where n is the size of our starting dataset). We are effectively selecting at random “with replacement”, which means that the same person can be selected twice, or indeed many times. (This is why I say that the population is effectively considered to have many copies of each person in it).

Therefore the boot strap data is the same size (same number of people in it) as the original dataset used to derive the logistic regression model. It is theoretically possible (though extremely unlikely) that a boot strap data set could be identical to that original dataset. However, it is far more likely that there will be differences, since some people will be included in the boot strap data set twice or more, and many are not included at all, although many would also be included just once.

Logistic regression of the same risk factors can then be applied to this boot strap sample, i.e. we rerun the logistic regression that gave us our chosen predictive model. However, we get slightly different regression coefficients, because of the modified sample. Prevalence estimates are then derived for each combination of risk factors, based on these new regression equations.

This process is repeated 1,000 times, to find 1,000 different boot strap samples, by random sampling processes, and to then fit logistic regression equations on each. The prevalence estimates are calculated for each combination of risk factors, for each of these 1,000 boot strap samples. For each region, a synthetic estimate is calculated for each boot strap sample, by appropriately weighting the prevalence estimates on each combination of risk factors (with the same weights as described above which reflect the anticipated prevalence of each combination of risk factors in the region). From these 1,000 synthetic estimates of prevalence of each region, a 95% confidence interval is calculated as the 2.5th to 97.5th centiles. Given that the estimates are distributed normally, these are taken to be mean +/- 1.96 SD (taking mean and SD of the 1,000 boot strap synthetic prevalence estimates for each specified region).

The following commands describe how this is done in Stata:

forvalues j=1/1000 {

use bootstrap, clear

(NB line above reads in original version of the data use used for logistic regrn eqn)

gen howmany=0

forvalues i=1/11516 {

local nn=floor(uniform()\*11516)+1

quietly replace howmany=howmany+1 if nnn==`nn'

}

nn is a random variable, derived from a uniform random variable which takes values between 0 and 1, to give a random variable between 1 and the total sample size.

The variable “howmany” records how many times each individual has been selected (for the specific bootstrap sample)

drop if howmany==0

expand howmany

The above 2 lines drop any people that have not been selected in our sample, and then repeat lines (twice or more) of any that have been selected twice or more.

quietly logit kneecategory2 agegp2 agegp3 palevelf1 palevelf2 palevelf3 smokef2 bmicatf22 bmicatf23 educ7 ssec8\_5 ssec8\_6 ssec8\_7 hobby1 [pweight=10\*probwtks]

The above lines run the chosen logistic regression on this boot strap sample of data, to get new estimates of regression coefficients.

use temp0, clear

The above reads in data set of all possible combination of risk factors, for purposes of calculating confidence intervals

\*\*\* the saved data set has 1 extra variable, so storing the extra bootstrapped estimate

predict est`j', xb

save temp0, replace

}

To get boot strap confidence intervals on specific regions, we need to firstly find predicted prevalences from these predicted log odds (by taking exp(log odds)/ [1+exp(log odds)] for each bootstrap estimate.

forvalues j=0/1000 {

gen prev`j'=exp(est`j')/(1+exp(est`j'))

}

Remember we are working on a data set with one line of data for each combination of risk factors. We then need to merge this data set, with the data set which gives appropriate weighted for each combination of risk factors for each region (which has many lines of data per region, 1152 for severe knee OA model).

merge agegp2 agegp3 palevelf1 palevelf2 palevelf3 smokef2 bmicatf22 bmicatf23 educ7 ssec8\_5 ssec8\_6 ssec8\_7 hobby1 using prevalences0

(This above commands lists each risk factor binary variable in the model as a variable that we are merging on).

For each boot strap sample, the synthetic prevalence estimate in any population is found by applying the same weights as above, according to the expected proportion of that population with any specified combination of risk factors (as follows – use of collapse command means that we conveniently end up with one line of data per informant).

forvalues j=0/1000 {

gen wt\_prev`j'=xyz\*prev`j'

}

collapse (sum) xyz wt\_prev\* (mean) c\_pt45p c\_tot\_mf\_ages, by(ccg\_code)

This gives 1,000 different synthetic estimates of prevalence for each population, one for each of the boot strapped samples of data. The confidence interval is found on these by taking the 2.5th and 97.5th centiles. Alternatively, the confidence interval can be found by taking the mean and SD of these prevalence estimates, and taking the mean +/- 1.96 SDs. [In practice, for estimates of severe knee OA, both these sets of estimates agreed very well, suggesting that the distribution of these estimates approximates very closely to the normal distribution – therefore the second method, using mean +/-1.96 SD, is a bit more precise]

egen meanr=rowmean(wt\_prev1-wt\_prev1000)

egen p2\_5r=rowpctile(wt\_prev1-wt\_prev1000), p(2.5)

egen p97\_5r=rowpctile(wt\_prev1-wt\_prev1000), p(97.5)

egen medianr=rowpctile(wt\_prev1-wt\_prev1000), p(50)

egen sdr=rowsd(wt\_prev1-wt\_prev1000)

Why is it not possible to put confidence intervals separately on each combination of risk factors? It is possible, but then averaging these would not agree to finding confidence intervals directly on appropriately weighted average prevalences of these, appropriate to specific populations. So that would not be a possible way forward with our objectives here.

Why is it necessary to divide the data into different groups by each combination of risk factors, rather simply taking account of the overall distribution of risk factors in the population? The weighted average of prevalences for a person with “average” risk factors is not the same as the weighted average prevalence, across all combinations of risk factors (appropriately weighted). The latter is what we want, and what we calculate directly.

This approach would work to find the appropriately weighted averaged log odds, since this a linear combination of risk factors. However, there is then a change of scale, taking the exponential to get the odds ratio, and then transforming again to get the prevalences.

1. Figures in brackets are 95% confidence intervals [↑](#footnote-ref-1)